NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

ASSESSING VULNERABILITIES IN INTERDEPENDENT INFRASTRUCTURES USING ATTACKER-DEFENDER MODELS

by

Cory A. Dixon

September 2011

Thesis Advisors:	David L. Alderson W. Matthew Carlyle
Second Reader:	Gerald G. Brown

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** September 2011	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE Assessing Vulnerabilities in Interdependent Infrastructures Using Attacker-Defender Models		**5. FUNDING NUMBERS**	
6. AUTHOR(S) Dixon, Cory A.			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School, Monterey, CA 93943-5000		**8. PERFORMING ORGANIZATION REPORT NUMBER**	
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A		**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**	
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number: N/A			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited		**12b. DISTRIBUTION CODE** A	

13. ABSTRACT (maximum 200 words)

Our economic and social welfare depend on certain "critical" infrastructures and key resources. Protecting these infrastructures is a challenge because they are complex, and as systems they are difficult to understand, predict and control. In addition, they do not operate in isolation, but are interdependent with other infrastructures. This presents a challenge for their modeling and analysis. Due to the complexity of modeling the operation of just a single infrastructure, most research to date has analyzed infrastructures in isolation. This thesis introduces a taxonomy of dependence relationships and incorporates these relationships into an attacker-defender model of interdependent infrastructure operation. We formulate and solve a sequence of models to illustrate how dependence relationships between infrastructures create vulnerabilities that are not apparent in single-infrastructure models, and we use the results to assess the consequences of disruptions to a system of infrastructures. We provide complete documentation for how to apply these techniques to real infrastructure problems and include a discussion of the necessary assumptions, as well as the pros and cons of our methods. Finally, we present examples of how to provide relevant, understandable results to help decision makers, such as where to make limited investments to increase resilience.

14. SUBJECT TERMS Infrastructure, Attacker-Defender, Dependence, Vulnerability, Operational Resilience, Interdiction, Worst-case Analysis			**15. NUMBER OF PAGES** 95
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500 Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

ASSESSING VULNERABILITIES IN INTERDEPENDENT INFRASTRUCTURES USING ATTACKER-DEFENDER MODELS

Cory A. Dixon
Commander, United States Navy
B.S., North Carolina State University, 1993

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN APPLIED SCIENCE
(OPERATIONS RESEARCH)

from the

NAVAL POSTGRADUATE SCHOOL
September 2011

Author: Cory A. Dixon

Approved by: David L. Alderson
Thesis Advisor

W. Matthew Carlyle
Thesis Advisor

Gerald G. Brown
Second Reader

Robert F. Dell
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Our economic and social welfare depend on certain “critical” infrastructures and key resources. Protecting these infrastructures is a challenge because they are complex, and as systems they are difficult to understand, predict and control. In addition, they do not operate in isolation, but are interdependent with other infrastructures. This presents a challenge for their modeling and analysis. Due to the complexity of modeling the operation of just a single infrastructure, most research to date has analyzed infrastructures in isolation. This thesis introduces a taxonomy of dependence relationships and incorporates these relationships into an attacker-defender model of interdependent infrastructure operation. We formulate and solve a sequence of models to illustrate how dependence relationships between infrastructures create vulnerabilities that are not apparent in single-infrastructure models, and we use the results to assess the consequences of disruptions to a system of infrastructures. We provide complete documentation for how to apply these techniques to real infrastructure problems and include a discussion of the necessary assumptions, as well as the pros and cons of our methods. Finally, we present examples of how to provide relevant, understandable results to help decision makers, such as where to make limited investments to increase resilience.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

Our economic and social welfare depend on certain "critical" infrastructures and key resources, such as energy, communication, and transportation systems. None of our nation's critical infrastructures operates in isolation. Each relies on inputs from other infrastructures to operate as intended, whether in the form of commodities, services, or information. These dependence relationships create potential vulnerabilities that are often not apparent until infrastructures are disrupted by accidents, failures, natural disasters, or deliberate attacks.

Because our nation cannot protect our critical infrastructures from all threats, we must assess their "resilience" in the face of disruption. Techniques for modeling infrastructure resilience vary, from risk-assessment models to "operational models" that attempt to capture component-level operational details and interactions. A natural starting point for assessing infrastructure resilience is to model each infrastructure in isolation. Modeling infrastructures individually allows us to accurately capture operational-level details appropriate to the particular type of system, and has encouraged the development of specific modeling techniques suited for the infrastructure at hand.

An attacker-defender (**AD**) model is a game-theoretic, operational model that assesses the worst-case disruption for an infrastructure operator by assuming component losses are selected by an intelligent attacker with perfect information. AD has been successfully applied to more than 150 case studies of individual infrastructures or military decision problems. However, in all cases, these analyses make the implicit assumption that any other supporting infrastructures are available and invulnerable to attack. This can result in inaccurate assessments of network resilience that provide operators with a false sense of security.

This thesis extends the standard attacker-defender model of a single infrastructure to account for the interdependence of two or more infrastructure systems. We present a general formulation for assessing resilience of a collection of independent infrastructures. We define a *direct*, cost-based dependence and introduce a model to examine such

relationships (e.g., *geographic* dependence). Finally, we define six *indirect* component-level dependence relationships: *single-input*, *exclusive-or*, *shared*, *substitute*, *complimentary* and *mutual*, and present a final formulation to assess the resilience of a collection of infrastructures containing both direct and indirect dependence relationships. We present an algorithm based on Benders decomposition to solve this formulation in an efficient manner.

To demonstrate our technique, we formulate and solve a sequence of simple network flow models and present the worst-case attacks and resulting operator flows for different levels of attacker resources. We show that disruptions are more costly when infrastructures are interdependent, and the presence of these dependence relationships favors the attacker. We show that locally optimal decisions of a single operator do not always lead to globally optimal behavior within a collection of interdependent infrastructures, necessitating the need for a decision maker to coordinate such activities at the global level.

We provide our formulations as a means of representing component-level dependence relationships in order to uncover resulting vulnerabilities and more accurately assess resilience for collections of infrastructures. Although we use minimum-cost network flow models in this thesis for ease of illustration, our main contribution does not depend on a network structure for the models used to represent the individual infrastructures. Natural extensions of our formulation include modeling of dependence classes other than physical, such as logical or cyber, along with the implementation of a tri-level (Defender-Attacker-Defender) model to identify an optimal defensive plan for the collection of infrastructures.

LIST OF ACRONYMS AND ABBREVIATIONS

AD	Attacker-Defender
CI/KR	Critical Infrastructures and Key Resources
DHS	Department of Homeland Security
DoT	Department of Transportation
ILN	Interdependent Layer Network
IRMF	Integrated Risk Management Framework
MILP	Mixed Integer Linear Program
NIAC	National Infrastructure Advisory Council
NPS	Naval Postgraduate School
PCCIP	President's Commission on Critical Infrastructure Protection

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGEMENTS

To my wife and best friend… Michelle. I could not have done this without your unwavering patience, continued understanding, and faithful support. I dedicate the rest of my life to making you happy.

To my parents, this would not have been possible without your enduring love, support and encouragement over the years. I thank you for allowing me to see the wisdom in education, and for giving me the drive to achieve my goals without fear of failure. I am forever in your debt.

To the best analyst and mentor I know, I thank my grandfather. You are the reason I have spent a career in the service of our great country. I thank you for providing more support and guidance during my formative years than you will ever now.

To my brother Phil, your immense drive and success are the example I strive to emulate. I am proud to be your twin.

Professor David L. Alderson, I thank you for the insight and advice you provided to mold this thesis into a cohesive body of work. Your ability to look beyond the direct contribution and tell the desired story is much appreciated, and your patience and tireless edits will not be forgotten.

Professor W. Matthew Carlyle, your positive attitude and motivation for this project were inspiring. Your analytic acumen proved immensely helpful, and I thank you for our many formulation and coding discussions.

Professor Gerald G. Brown, I thank you for lending your time and experience to this work. Your extensive knowledge of the subject matter provided the necessary insight that was crucial to the completion of this thesis.

Last, but certainly not least, to my classmates. I thank you for the collaboration and the team-first approach we maintained throughout this academic endeavor. I am richer for having known each and every one of you.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTERDEPENDENT INFRASTRUCTURE SYSTEMS

A. INTRODUCTION

The September 11, 2001, attacks by Al Qaeda on the World Trade Center caused over 240 disruptions across eight infrastructures within Manhattan, including over 50 identified as resulting from interdependence relationships (Wallace, Mendonca, Lee, Mitchell, & Chow, 2003). When considered in concert with the successful attack on the Pentagon and the failed attempt on the United States Capitol building on that day, these incidents had short-term effects on global financial markets and still impact worldwide air transportation today.

Society's economic and social welfare depend on certain "critical" infrastructures and key resources, such as energy, communication, and transportation systems. The U.S. Department of Homeland Security (DHS) explicitly lists eighteen critical U.S. infrastructure and key resource (CI/KR) sectors vital to our nation's security in the National Infrastructure Protection Plan (DHS, 2009). In the 2007 National Strategy for Homeland Security, DHS recognizes it is not possible to deter all threats to our infrastructure; thus, the nation must mitigate vulnerabilities by "ensuring the structural and operational resilience" of CI/KR (p. 27). The National Infrastructure Advisory Council (NIAC) defines resilience in the following manner:

> Infrastructure resilience is the ability to reduce the magnitude and/or duration of disruptive events. The effectiveness of a resilient infrastructure or enterprise depends upon its ability to anticipate, absorb, adapt to, and/or rapidly recover from a potentially disruptive event. (2009, p. 8)

None of our nation's critical infrastructures operates in isolation. Each relies on inputs from other infrastructures to operate as intended, whether inputs in the form of commodities, services, or information. Even when considered in isolation, these infrastructures can be complex and sizable when viewed at an operational level. A regional power grid, for example, might contain thousands of power lines and buses, and hundreds of generators (Salmerón, Wood, & Baldick, 2009). Infrastructures are also continually changing to meet new demands or exploit new technology. Protecting our

infrastructures is a challenge because they are complex and difficult to understand, predict and control. This presents a challenge for their modeling and analysis.

Techniques for modeling critical infrastructure resilience vary, from risk assessment models to "operational models" that attempt to capture component-level operational details and interactions. Because of the details required to capture the operation of just a single infrastructure, most researchers to date have modeled infrastructures in isolation. In a few cases, researchers have formulated models of *collections* of infrastructures and their interdependence relationships to assess vulnerabilities. To the best of our knowledge, however, no one has modeled a collection of infrastructures with the level of fidelity necessary to assess operational resilience.

B. RELATED WORK

We first discuss a brief history of critical infrastructure protection and then introduce a baseline model of infrastructure operation in isolation. We then summarize how others define, group and model interdependence relationships. Finally, we present our contribution in context.

1. Critical Infrastructure Protection

In *Critical Path: A Brief History of Critical Infrastructure Protection in the United States*, Brown (2006) documents the rise of critical infrastructure in the United States from the earliest postal road networks in the late eighteenth century to present day. The basis for modern discussion on critical infrastructure originated as a result of late twentieth-century events, specifically the dramatic rise in terrorist activity against the United States in the early 1990s: the bombing of the World Trade Center in 1993, the Khobar Towers in Saudi Arabia in 1993, and a federal building in Oklahoma City in 1995, for example. Subsequently, President Clinton established the President's Commission on Critical Infrastructure Protection (PCCIP) through Executive Order No. 13010 (1996).

> It is clear to us that infrastructure assurance must be a high priority for the nation in the Information Age. With escalating dependence on information and telecommunications, our infrastructures no longer enjoy the protection of oceans and military forces. They are vulnerable in new ways. We must protect them in new ways. (PCCIP, 1997b, p. 9)

In its report *Critical Foundations: Protecting America's Infrastructures* the PCCIP recognizes that the U.S. suffers an increasing dependence on critical infrastructure while systemic vulnerabilities grow, due to new cyber threats and increasing system complexities and interdependence relationships. The report also recognizes a wide spectrum of threats, notes a deficient general awareness by the public, and expresses concern over a lack of a national focus. As a result, the PCCIP concludes that the increasing threats in new domains require new thinking from both the public and private sectors, with immediate action needed to protect our future. PCCIP's recommendations include establishing a national organization to include CI/KR sector coordinators, agencies and government councils to bridge the gap between public and private sectors while enhancing information sharing and cooperation, revising outdated regulations to account for changes in technology; and changing research and development goals to counter current sector weaknesses (PCCIP, 1997b).

The United States is exposed to escalating hazards present as a consequence of infrastructure interdependence brought about by both technology and the Internet (PCCIP, 1997a; Brown, 2006). As our sophisticated networks increasingly depend on computers and Internet connections to automate many of their routine functions, infrastructure interdependence grows. We often realize these interdependence relationships only when bad things happen. A notable example is the Northeast Blackout of 2003, the worst in U.S. history. A failure of several transmission lines in Ohio cascaded into power failures across the Northeast United States and Canada, leaving an estimated 50 million people without power (Davidson, 2008). It also resulted in loss of water supply due to inadequate pumping, regional transportation outages due to rail and airline stoppages, and temporary interruption of cellular telephone service.

Whether threatened by human error, natural disaster, or terrorists, our nation's critical infrastructures are at risk. Their protection requires methods of identifying and modeling interdependence relationships, detecting vulnerabilities and placing defenses to protect the most critical assets.

2. Operational versus Nonoperational Analysis

The National Strategy for Homeland Security recognizes the need for "operational resilience" in our infrastructure systems (DHS 2007, p. 27). What does this mean? According to the NIAC (2009), it is the ability of an infrastructure, when faced with disruptions, to adjust its activities and continue functioning (or quickly recover) to meet its objective.

To assess operational resilience, we must model the function of the infrastructure. We cannot do this by representing only the various system components and assessing their individual vulnerabilities. We need to represent how the various components work together to accomplish the infrastructure purpose, or objective (i.e., function). We must be able to recognize how operations change as a result of infrastructure activity decisions. Decisions can follow rule sets (e.g., if component **A** fails, switch control to component **B**), but they often require us to evaluate tradeoffs that result from unexpected consequences after some disruption. Thus, in order to assess infrastructure operational resilience, we must capture the component-level details and the interaction between the components as decisions are made. We refer to a model that captures this level of fidelity as an *operational* model.

We consider any model that does not capture the operation of an infrastructure to be *non-operational*. An example is the Integrated Risk Management Framework (IRMF) advocated by DHS in its National Infrastructure Protection Plan (DHS, 2009). This framework assesses infrastructure resilience through risk analysis as a function of threats, vulnerabilities, and consequences. It does not capture the operational details of infrastructure. We prefer the more prescriptive nature of operational models and focus our discussion on these here.

3. Modeling Infrastructures in Isolation

A natural starting point for assessing infrastructure resilience is to model each infrastructure in isolation. Modeling infrastructures individually allows us to accurately capture operational-level details appropriate to the particular type of system. This has encouraged the development of specific modeling techniques suited to the infrastructure at hand. Regardless of technique used, the goal is to assess infrastructure performance in the presence of (possibly uncertain) disruption.

a. Single Scenario Performance Analysis

The basis for our analysis is a mathematical formulation representing the operation of an infrastructure. Brown, Carlyle, Salmerón and Wood (2005, 2006) catalogue various ways of representing infrastructure operation. In many cases, a linear representation with an objective function based upon cost is sufficient to capture first-order effects.

Without loss of generality, an "operator" chooses a set of activities Y to minimize system operating cost, subject to specific infrastructure constraints and known, fixed disruptions–such as those caused by system component malfunctions, acts of nature, or known attacks. The resulting model is

$$\min_{Y} f(Y, \overline{X}). \tag{1.1}$$

Here, $\overline{X}$ represents these known, fixed component-level disruptions. A key assumption is that even with known disruptions, the operator will control his infrastructure, by choosing or rewarding actions Y, to minimize total cost. By choosing an appropriate $\overline{X}$ for any scenario of interest, we can conduct systematic "what-if" analyses. The function f represents the total system operating cost.

b. Expected Performance Analysis

In many cases, it is appropriate to characterize the disruption or loss of system components using probabilities (e.g., a weather event or engineering failure).

Given a random variable $\tilde{X}$ representing the availability of system components, we can define the *expected performance* of the system as

$$E_{\tilde{X}}\left(\min_{Y} f(Y, \tilde{X})\right). \tag{1.2}$$

Here, for any fixed realization $\overline{X}$ of the random variable $\tilde{X}$, the system operator chooses the best activities as in Equation (1.1). We can evaluate this expectation (1.2) using traditional Monte Carlo techniques (Law & Kelton, 2000, pp. 90-91).

c. Worst-Case Performance Analysis

When we assume disruptions are selected by an intelligent attacker with perfect information, the model becomes a two-stage, zero-sum game if we assume the attacker wishes to maximize what the operator seeks to minimize. It is difficult to keep details of critical infrastructures hidden with absolute certainty, so assuming the attacker has the information he needs is conservative, but prudent, and it sets a worst-case for the operator, or "defender" (Brown et al., 2006). It can also be used as a model of "insider threat," or "competitor threat." In either case, we also assume the attacker and defender actions must be sequential, with the attacker first choosing disruptions (X) to the network to maximize the defender's operating costs, followed by the defender choosing a set of activities (Y) to minimize the resulting operating costs. The resulting model is

$$\max_{X} \min_{Y} f(Y, X).$$

This attacker-defender (**AD**) model has been successfully applied in the analysis of many infrastructures over the past decade: electric grids (Salmerón, Wood, & Baldick, 2004, 2009), theater ballistic missile defense (Brown, Carlyle, Diehl, Kline, & Wood, 2005), oil pipelines and airport security (Brown et al., 2005, 2006), and transportation systems (Alderson, Brown, Carlyle, & Wood, 2011). There have been more than 150 Naval Postgraduate School (NPS) case studies of individual infrastructures or military decision problems that have used **AD** models.

*All of these prior analyses with **AD** models consider only infrastructures in isolation.*

Each of the above techniques identifies sets of critical components whose defense improves resilience within an individual infrastructure. In all cases, these analyses also make the implicit assumption that any other supporting infrastructures are available and invulnerable to attack. This can result in inaccurate assessments of network resilience that provide operators with a false sense of security.

For example, consider the work of Salmerón et al. (2004, 2009) modeling electric power grids. They formulate and solve AD models to help utility companies identify critical components and improve grid resilience. While Salmerón et al. analyze vulnerabilities to physical attacks on the grid in detail, they acknowledge their work does not account for cyber attacks. An electric grid's reliance on telecommunications and Internet for control and on coal, natural gas, petroleum or water for generation is a significant vulnerability.

4. Representing Dependence

Rinaldi, Peerenboom and Kelly (2001) define interdependence as "a bidirectional relationship between two infrastructures through which the state of each infrastructure influences or is correlated to the state of the other" (p. 14). While it is perhaps obvious that infrastructure dependence relationships are crucial, it is less clear how to identify and subsequently represent them appropriately in an analysis. Chou and Tseng (2010) describe a technique for automated "knowledge discovery" of critical infrastructure interdependence relationships based on the use of data mining in sets of recorded infrastructure failure data (p. 539). Other researchers have proposed categories for dependence relationships, specifically based on *class* and *type*.

a. Class

It is sometimes convenient to categorize infrastructure interdependence relationships by *class*. Rinaldi et al. (2001) offer four classes as defined in previous literature: *physical*, *cyber*, *geographic* and *logical*. Using their definitions, a *physical* dependence is based upon the physical flow of a commodity (e.g., oil needed to run an electric generator), while *cyber* depends upon some form of transmitted data (e.g., supervisory control and data acquisition control of a steam valve). Similarly, a

geographic dependence exists when components of separate infrastructures are geographically co-located, such that a physical attack or disruption to one system impacts other systems. For example, consider the attack on a railroad bridge that also serves as a supporting structure for electrical and telecommunications cables. It has been noted that a geographic relationship between components is a correlation, and not a true dependence (Bernstein, Bienstock, Hay, Uzunoglu, & Zussman, 2011). Nevertheless, to keep in line with the bulk of literature, we refer to a geographic relationship here as a dependence. Rinaldi et al. (2001) place all other dependence relationships that are not physical, cyber or geographic into the *logical* class, such as those driven by policy, contractual or legal obligations, or market forces.

b. Type

In addition to categorizing dependence by class, it is sometimes helpful to define these relationships by *type*. We consider five types used by other researchers: *input, shared, exclusive-or, mutual,* and *co-located* (Wallace et al., 2003; Lee, Mitchell, & Wallace, 2004, 2007). For purposes of explanation, consider two infrastructures, **A** and **B**. An *input* dependence is a one-way relationship between infrastructures; system **A** receives input from system **B**, but system **B** receives no input from system **A**. A *shared* dependence exists when systems **A** and **B** share a common component or service. An *exclusive-or* dependence arises when either system **A** or system **B**, but not both, can use a component or service. A *mutual* dependence between infrastructures is a two-way relationship where both systems **A** and **B** receive input from each other, although not usually from interactions of same components or services. Finally, in a *co-located* relationship, the systems do not depend on each other, but are located within the same geographic region, and may be similarly affected by some event (e.g., a natural disaster). The *co-located* dependence type is equivalent to the *geographic* class defined by Rinaldi et al. (2001). Our models expand directly from these dependence types.

5. Operational Models with Dependence

Although we use minimum-cost network flow models to represent the operation of individual infrastructure systems in this thesis, in a real application we might need

something more complicated. For example, electrical transmission models represent the physics of power flow, and these are not amenable to a simple network flow model. However, our main contribution is independent of the particular models used to represent the individual infrastructures. What is important is the way that we model the interdependence relationships *between* infrastructure systems, and we do so using notation and terminology derived from the study of network flows. We represent each of the various *classes* of relationships between infrastructure systems by modeling the flow of a commodity (e.g., electric power) from one system to another and by capturing how a sufficient flow of that commodity enables an activity (e.g., operating a particular water pump) in the other system.

Before introducing our work, we examine the contributions of Kennedy (2009) and Lee et al. (2007). They use *network design* models to demonstrate *geographic* and *physical* dependence respectively within collections of infrastructures.

a. Layered Interdependent Infrastructures

Kennedy (2009) models network dependence between multiple infrastructures in a single network flow model with two sets of variables. The first set of variables represents individual infrastructure activities, while the second, a set of binary variables, represents the satisfaction of dependence relationship requirements between infrastructures using binary directed arcs. Each infrastructure exists as a separate layer within the model, connected to others through the dependence relationships (arcs) that exist between distinguished interdependent nodes. He considers this collection of infrastructure layers and dependence relationships as a single, unified, directed graph of all flow activities, expressed as a single commodity.

Kennedy defines “effects options” that can influence a subset of the interdependent components (nodes and arcs) within the collection of layered networks, and for each effect option, he defines the change on the individual components produced by the effect. He uses a cost-based formulation composed of two parts: the cost (or benefit) for activities within each layer of the network, and the cost (or benefit) of each effect option. Each node has an associated supply or demand, and each arc has a

maximum capacity. Kennedy then models the layered network as a minimum-cost network and solves this monolithic model using Benders decomposition.

Kennedy presents a reformulation of the classic network design problem with notation that reveals its applicability to design problems on layered networks. His models are one-sided and do not involve an intelligent adversary. For example, his formulation may be an attacker's problem with the objective to either inflict maximum damage or fully disrupt a collection of infrastructures for a minimum cost. The attacker's costs would consist of both the costs to disrupt flow within each infrastructure and the costs associated with a particular effect option to cut selected interdependent arcs in the collection. Similarly, in an operator's problem, the costs might represent activity costs within each infrastructure, along with upgrade option costs for the collection. Kennedy solves by enumerating through the finite number of effects options and interdependent links to identify the combination in this searched set that produces the desired result.

The examples used to demonstrate Kennedy's model assume a geographic (or co-located) source of dependence, for which his method is a viable solution. However, it is not clear how to apply Kennedy's techniques to any dependence other than geographic co-location. In order to model physical, cyber or logical links, one must define a more robust formulation than the binary link between infrastructure layers used in his work.

b. Mixed-Integer Network Flow Model

Lee et al. (2007) define mathematics to explicitly model the five types of infrastructure dependence relationships discussed in their previous work (Wallace et al., 2003). They show the modeling value with a case study. They argue that prior efforts to represent dependence generally fall short because those efforts develop hybrid models that do not adequately capture the operation of the infrastructures, and therefore lose their value to the individual infrastructure operators. Lee et al. create an interdependent layer network (ILN) of the electric, subway and telecommunications networks in lower Manhattan, model a disruption to the ILN and demonstrate the model's usefulness in restoration of services.

We summarize the modeling methods introduced by Lee et al. (2007) for each of the five types of dependence relationships below.

(1) *Input Dependence*. When an infrastructure *h* requires input from infrastructure *i*, Lee et al. (2007) model the location at which this dependence occurs as a transshipment node *l* in *h*, and use a binary *connector variable* $y_{h,l}^{i,j}$ to represent adequate supply at node *j* in *i* to allow the function of node *l* in *h* (Figure 1).

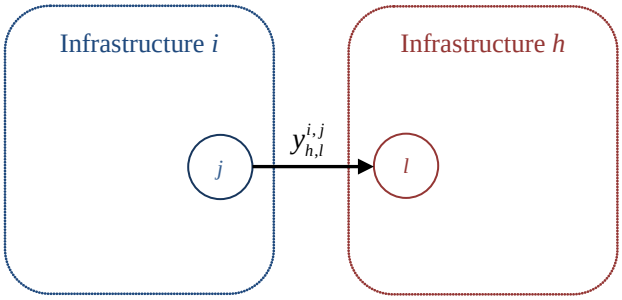

Figure 1. *Input* Dependence according to Lee et al.

Node *l* in infrastructure *h* requires input from node *j* in infrastructure *i*. The binary connector variable $y_{h,l}^{i,j}$ represents whether or not adequate supply is provided.

Lee et al. (2007) use balance-of-flow equations to determine the potential commodity *shortfall*, s_j^i, at any node *j* in infrastructure *i*. They also give each node a weighting factor k_j^i, and, while not explicitly stated, we assume this weighting factor is a penalty assigned for not meeting demand. The ILN objective function contains a term that is the product of each shortfall variable with its respective weighting factor

$$\sum_{i \in I} \sum_{j \in J} k_j^i s_j^i .$$

Lee et al. (2007) include a constraint that relates the shortfall variable to the connector variable, ensuring the linking variable allows operation of node *l* $\left(y_{h,l}^{i,j} = 1\right)$ only if no shortfall at node *i* exists, i.e.,

$$s_j^i \leq (1 - y_{h,l}^{i,j})(-b_j^i),$$

for certain combinations of *i*, *j*, *h*, and *l*.

Defining the relationship between the operational capacity of node l and the state of the connector variable requires additional constraints, and each is involved via a linear relationship. For example, if l is a transshipment node, its flow capacity is the product of its rated capacity w_l^h and the connector variable $y_{h,l}^{i,j}$ as shown in the following equation (note $e \in \delta^+(l)$ represents all inbound arcs to node l and x_e^h is arc flow in infrastructure h):

$$\sum_{e \in \delta^+(l)} x_e^h \leq w_l^h y_{h,l}^{i,j},$$

again, for certain combinations of i, j, h and l.

Therefore, if the connector variable is zero, then the node capacity is also zero, but when the node connector variable is one, the node capacity equals its rated capacity. Lee et al. discuss the possibility that the demand node's capacity can vary in other ways, as opposed to simply switching between the rated capacity and zero. Their modeling can accommodate this.

Although the Lee et al. formulation is more complex than the simplification described above (with differing subsets and constraints for demand, supply and transshipment nodes), the main concept remains the same. Commodity shortfall at a parent node that supplies commodity to another infrastructure prevents operation of that supported child node. In addition, weighted slack variables are included in the objective function. Minimizing the overall objective function tends to drive all slack variables to zero, thereby supplying the necessary flow between interdependent networks, if possible.

(2) *Mutual Dependence*. Lee et al. (2007) consider mutual dependence in terms of infrastructures (**A** and **B**), where infrastructure **A** relies on infrastructure **B** for some supply and vice versa. In this case, a mutual dependence becomes two sets of input dependence relationships, where a parent node in **A** supports a child node in **B**, and a separate parent node in **B** supports a child node in **A**. Modeling then follows that of the *input* dependence.

(3) *Shared Dependence*. Lee et al. (2007) consider shared dependence to represent multiple commodities flowing across the same arcs or nodes.

They, therefore, define $n \in N$ as a commodity in the set of commodities *N*, and subsequently ensure the sum of all commodity flows across an arc or node is less than the component capacity

$$\sum_{n \in N} x_{e,n}^{i} \leq u_{e}^{i}, \tag{1.3}$$

for each arc *e* of infrastructure *i*.

(4) *Exclusive-or Dependence*. This dependence is similar to shared dependence where a component is available to two or more commodities, but only one can use it at a time. This requires the addition of a commodity flag (binary) to Equation (1.3) and a subsequent constraint restricting the number of commodities as follows:

$$x_{e,n}^{i} \leq u_{e}^{i} r_{e,n}^{i} \tag{1.4}$$

$$\sum_{n \in N} r_{e,n}^{i} \leq 1, \tag{1.5}$$

again, for each arc *e* of infrastructure *i*.

Equation (1.4) serves the same purpose as (1.3) where *r* turns the flow of *i* on arc *e* on or off as needed. Equation (2.5) restricts the number of commodities that can flow on arc *e* simultaneously (here, just one).

(5) *Co-located Dependence*. Lee et al. (2007) do not explicitly consider co-location within their formulation, but leave it to the operator to determine co-location effects and manually adjust the model to account for the changes in supply, demand or capacity.

To summarize, the ILN formulation and case study introduced by Lee et al. (2007) show it is possible to represent infrastructure dependence relationships in a unified model without losing required fidelity of any particular system model.

While they discuss *shared* and *exclusive-or* dependence in regards to sharing flow of multiple commodities across a common node or arc (pipelines or roads), we argue this description does not fully represent the possibilities of either type of dependence. Instead, we model *shared* and *exclusive-or* dependence in terms of what is

required for operation of a child node, as opposed to simply the restrictions on type of commodity flow within a single infrastructure. Lee et al.'s formulation is designed to minimize the cost to restore services post-attack, and it does not consider an intelligent attacker; thus, it serves as another example of a network design model.

6. Our Contribution in Context

We use the attacker-defender model as a basis for a worst-case analysis, which is "crucial to a credible assessment of infrastructure vulnerability and for planning mitigating actions" (Brown et al., 2006, p. 543). We take advantage of both the cost-based, co-location formulation of Kennedy (2009) and dependence *type* formulations of Lee et al. (2007), and introduce a taxonomy of dependence relationships. We incorporate each dependence type into an operational-level **AD** model of infrastructure behavior, we formulate and solve a sequence of models to illustrate how the dependence relationships create vulnerabilities that are not apparent in the single-infrastructure models, and we assess the consequences of disruptions to the system of infrastructures. We provide complete documentation for how to apply these techniques to real infrastructure problems and include a discussion of the necessary assumptions, as well as the pros and cons of our methods. Finally, we present examples of how to provide relevant, understandable results to help decision makers, such as where to add limited investments.

II. MODEL FORMULATION

A. SINGLE INFRASTRUCTURE IN ISOLATION

We begin with a model for the operation of a single infrastructure in isolation. Without loss of generality, we assume the objective is to minimize the cost associated with infrastructure operation. In what follows, we will refer to the operator as the "defender' to keep in line with the literature. We define necessary terms and present the basic defender's problem before introducing the concept of an attacker and formulating the attacker's problem.

1. Defender Problem (D)

For clarity of exposition, we formulate each defender's problem as an optimization of commodity flows over a set of nodes and arcs. (We stress again that much more general models are admitted by our methods, but these networks are easy to illustrate and discuss.) These models involve balance of flow constraints and can have additional side constraints on the flow. We adopt standard definitions and notation common to the study of network flows, formally defined in Ahuja, Magnanti and Orlin (1993); see Figure 2. Let $g=(N,A)$ represent a directed graph, where N is a set of nodes and $A \subseteq N \times N$ is a set of directed arcs. For each arc $(i, j) \in A$, we let Y_{ij} denote commodity flow from i to j, u_{ij} represent the arc capacity, and c_{ij} indicate the normal operating cost per unit of commodity flow across arc $(i, j) \in A$. For each $n \in N$, we define the commodity supply as b_n, with $b_n>0$ for supply and $b_n<0$ for demand. We refer to a node that has no supply or demand (i.e., $b_n=0$) as a transshipment node.

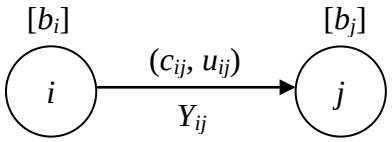

Figure 2. Commodity flow representation along directed arc $(i, j) \in A$.

Infrastructure operators are often concerned with both excess commodity and shortages within their network as both may incur network costs and even degrade infrastructure function. For this reason, we define an excess ($EXCESS_n$) and shortage ($SHORT_n$) of commodity per node and associated per-unit costs associated with each, $ePen_n$ and $aPen_n$ respectively (these costs can be zero if, for example, excesses are acceptable). When the balance of flow constraint for node *n* is formulated using these variables (and associated costs), we say that this constraint has been made *elastic*, and call these additional variables that measure conventional constraint violations *elastic variables*.

No infrastructure is immune to disruption, whether a pump seizes in a water system, a power line in an electric grid snaps from a winter storm, or an enemy bombs a bridge in a transportation network. We collectively refer to any component loss in our models as an "interdiction," where it can be caused by a mechanical failure, a random act of nature, the deliberate actions of an intelligent adversary, or a host of other reasons. We consider two types: arc interdiction and node interdiction. In our model, we let the binary value $\bar{X}_{ij}$ represent the interdiction of arc $(i, j) \in A$, with $\bar{X}_{ij} = 0$ representing a fully functioning arc and $\bar{X}_{ij} = 1$ indicating the arc has been interdicted. In this model, an interdicted arc $(i, j) \in A$ has an additional per-unit operating cost q_{ij}. Presumably, this cost can be high enough to preclude any function of the interdicted arc at all.

We also wish to model node interdiction. We do this through "node-splitting," whereby a node *n* is replicated into two nodes, *n'* and *n"* with a single directed arc $(n', n") \in A$ allowing flow between them. All inbound arcs (*i*, *n*) to *n* for $i \in N$ enter *n'* and all outbound arcs (*n*, *j*) from *n* for $j \in N$ exit *n"* (Figure 3). With this modification, interdiction of directed arc (*n'*, *n"*) is equivalent to interdiction of the original node *n*. Note that the interdicted cost for node *n* is $q_{n'n"}$ (after node-splitting), and that presumably this cost can be high enough to preclude any flow.

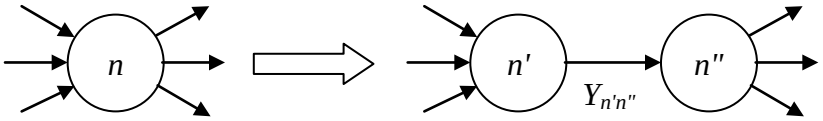

Figure 3. Representation of node-splitting.

The cost to operate an arc is a product of the amount of flow on an arc (Y_{ij}) and the total per-unit operating cost for that arc. This total arc cost consists of the normal per-unit operating cost (c_{ij}), and, if the arc is not functional $\left(\bar{X}_{ij} = 1\right)$, the per-unit interdiction cost (q_{ij}). The total operating cost for the network is the sum of the operating costs across the set A of all arcs in the network.

This per-unit cost q_{ij} is equivalent to paying a premium to deliver flow, possibly by going outside the system. For example, shipping oil along an interdicted pipe segment could mean hiring a caravan of trucks to move the flow around the unusable section of pipe. This could be very expensive, but if the alternatives are worse, we expect that an operator will do it. By setting q_{ij} carefully, we ensure the model reflects the desired behavior.

We represent the penalty cost at a node by the product of its commodity shortage ($SHORT_n$) and its penalty cost per unit ($sPen_n$) or the product of its excess supply amount ($EXCESS_n$) and its penalty cost ($ePen_n$). The total penalty cost for the infrastructure equals the sum of all penalty costs over the set of nodes, N.

We summarize the single-infrastructure defender problem **(D)** as follows.

Indices and index sets [~cardinality]

$n \in N$	Nodes in an infrastructure	(alias i, j)	[moderate]
$(i, j) \in A \subseteq N \times N$	Directed arcs (edges with direction included)		[many]

Given Data [units]

b_n	Supply (a negative value indicates a demand) of commodity at node n	[commodity unit]
$ePen_n$	Penalty for commodity excess at node n	[cost/unit]
$sPen_n$	Penalty for commodity shortage at node n	[cost/unit]
u_{ij}	Capacity of arc $(i, j) \in A$	[commodity unit]
c_{ij}	Cost per-unit of operating arc $(i, j) \in A$	[cost/unit]
q_{ij}	Additional cost per-unit of operating interdicted arc $(i, j) \in A$	[cost/unit]
$\bar{X}_{ij}$	Indicates interdiction of arc $(i, j) \in A$ $\bar{X}_{ij} = 1$ is interdicted, $\bar{X}_{ij} = 0$ not interdicted	[binary]

Decision variables [units]

$EXCESS_n$	Commodity excess at node n	[commodity unit]
$SHORT_n$	Commodity shortage at node n	[commodity unit]
Y_{ij}	Flow on $(i, j) \in A$	[commodity unit]

Formulation (D)

$$\min_{\substack{Y \\ EXCESS,\, SHORT}} \; f(Y, \bar{X}) \equiv \sum_{(i,j)\in A} \left(c_{ij} + q_{ij}\bar{X}_{ij}\right) Y_{ij} + \sum_{n\in N} \left(ePen_n EXCESS_n + sPen_n SHORT_n\right) \tag{2.1}$$

subject to :

$$\sum_{j:(n,j)\in A} Y_{nj} - \sum_{i:(i,n)\in A} Y_{in} + EXCESS_n - SHORT_n = b_n \qquad \forall n \in N \tag{2.2}$$

$$0 \le Y_{ij} \le u_{ij} \qquad \forall (i, j) \in A \tag{2.3}$$

$$EXCESS_n \ge 0,\; SHORT_n \ge 0 \qquad \forall n \in N \tag{2.4}$$

$$Y \in \Psi \tag{2.5}$$

This single-commodity minimum cost network flow model represents the defender's problem. Given fixed disruptions $\overline{X}$, the objective function (2.1) expresses the operating cost incurred by commodity flow *Y*, while constraint (2.2) represents the balance of flow for every node in the network. We restrict commodity flow (Y_{ij}) between the arc capacity (u_{ij}) and zero in constraint (2.3). Equation (2.4) is the non-negativity constraint for the elastic variables, and Equation (2.5) represents any other side constraints pertaining to the activities *Y*, whether physical or logical. While Equation (2.1) explicitly shows the minimization is with regard to variables *Y*, *EXCESS* and *SHORT*, we will subsequently write "$\min_{Y}$" for brevity with the understanding that we are also minimizing with respect to the *EXCESS* and *SHORT* variables, as well.

The infrastructure operator is myopic in this problem; he does not concern himself with anything outside his individual infrastructure. He observes, and may actually set as constants exogenous to his optimization, the interdicted and non-interdicted arc costs, node supply and demand, and penalties for commodity shortages or excess supply. In addition, even though the operator may not control the functionality of system components, the formulation allows him to adapt to identified outages $\left(\overline{X}_{ij} = 1\right)$ and operate his degraded system so as to minimize cost.

2. Attacker Problem (AD)

Consider an intelligent adversary (attacker) wishing to disrupt the system operation. Brown et al. (2005) show that several assumptions are required for this problem. First, we assume the defender will always operate his infrastructure optimally. Regardless of attacks or component failures, the defender will adjust his system operation as necessary to continue to minimize his cost of operation. Therefore, for a given level of attack resources (Γ), we assume the attacker's goal is to maximize the defender's resulting cost. X_{ij} is now a decision variable for the attacker. We also assume the attacker has perfect knowledge of the targeted infrastructure. This may not be true, but conservatively allows us to identify the worst-case scenario for the defender. The attacker plans his attacks to maximize the defender's minimum cost given his attack

resources, knowing that the defender will subsequently adjust his operations as required to minimize cost.

We summarize the single-infrastructure attacker problem **(AD)** as follows.

Decision variables [units]

X_{ij} Indicates interdiction of arc $(i, j) \in A$ [binary]

$X_{ij} = 1$ is interdicted, $X_{ij} = 0$ not interdicted

Formulation (AD)

$$\max_{X} \min_{Y} f(Y, X) \equiv \sum_{(i,j)\in A} \left(c_{ij} + q_{ij} X_{ij}\right) Y_{ij} + \sum_{n \in N} \left(ePen_n EXCESS_n + sPen_n SHORT_n\right) \quad (2.6)$$

subject to:

$$\sum_{j:(n,j)\in A} Y_{nj} - \sum_{i:(i,n)\in A} Y_{in} + EXCESS_n - SHORT_n = b_n \qquad \forall n \in N \quad (2.2)$$

$$0 \le Y_{ij} \le u_{ij} \qquad \forall (i, j) \in A \quad (2.3)$$

$$EXCESS_n \ge 0,\ SHORT_n \ge 0 \qquad \forall n \in N \quad (2.4)$$

$$Y \in \Psi \quad (2.5)$$

$$X \in \Gamma \quad (2.7)$$

By adding the attacker to the original defender's problem (**D**), the new formulation (**AD**) becomes a two-stage optimization problem. The objective function (2.6) expresses with respect to attacks (X) the operating cost incurred by commodity flow (Y). Constraints (2.2–2.5) remain the same as in the defender's problem. Equation (2.7) restricts the attack resources to a given feasible region, Γ. As stated here, attacks (X) interdict directed arcs. In cases where an attack impedes flow in both directions (e.g., both sides of a divided highway), we can additionally require that $X_{ij} = X_{ji}$.

The importance of **AD** lies in the identification of the most vulnerable components, or "critical assets" of an infrastructure for a given attack level. This has led to much work finding various infrastructure vulnerabilities and mitigating them, championed by researchers at the Naval Postgraduate School (Salmerón et al., 2004; Brown et al., 2005, 2006; Brown, Carlyle, & Wood, 2008, Alderson et al., 2011). Model **AD** serves as the building block for our infrastructure interdependence models.

B. MULTIPLE INDEPENDENT OPERATORS

We now consider a more global perspective, one in which we have a collection of infrastructures. We assume that each infrastructure is independent and operated by a separate system operator concerned with minimizing his individual infrastructure operating costs. As in the single-infrastructure case, each operator disregards the surrounding infrastructures. We introduce the concept of a *global manager* wishing to minimize the cost of the entire collection of infrastructures, we define necessary terms, and we present the basic problem structure before introducing the concept of an adversary and subsequent formulation of the attacker problem. The global manager is now the "defender" as opposed to any of the independent selfish infrastructure operators.

1. Defender Problem (MULTI-D)

Consider a set of infrastructures R, where $r \in R$ represents an individual infrastructure. We assume all nodes are unique and that each is present in only one infrastructure. Therefore, let $R(n) \in R$ denote the infrastructure in which node n resides, while $N_r \subseteq N$ denotes the set of all nodes in infrastructure r.

How does a global manager of a collection of infrastructures value its performance? To answer this, we assume the manager desires to minimize the total operating cost of the collection, just as each individual operator seeks to minimize the cost of his respective infrastructure. Therefore, the global manager's objective function is built in the same manner, with costs tied to everyday operation and penalties for unmet requirements.

Individual infrastructure operators may measure their operating costs in different ways. The owner of an electric grid might think in terms of megawatt hours (MWh) of electricity, while the oil producer might think in terms of barrels of oil. A global manager might use current prices ($/MWh and $/barrel) to convert to a uniform standard, dollars ($). We introduce h^r to serve as this relative *cost conversion factor* for each infrastructure, and multiply it by the network operating costs of infrastructure r.

The global manager must also specify a way to compare the penalty costs for different infrastructures. However, unlike operating costs, the penalty costs are more subjective. For example, consider the operators of a city water infrastructure and a local metro system. Within their respective infrastructures, each owner might individually impose stiff penalties for not meeting demand. Nevertheless, in the event of a major disruption, a city manager overseeing both infrastructures might value water distribution to the citizens as a higher priority than mass transit because of secondary effects (dehydration, sickness) that come from not having adequate potable water. In this case, the global penalties for water shortages could reflect this and be set higher than those for shortages in the mass transit system. We implement this in our model through use of p^r, a *policy weighting* given to each infrastructure. Like h^r, we must ensure that any policy weights introduced yield objective function values in standard units (e.g., dollars).

The parameter h^r is a way to equalize costs; it can be set by individual operators if a common cost baseline has been set by the global manager of the infrastructure collection. Setting the system-wide policy weights (p^r) requires either a quantitative assessment of secondary effects or a qualitative assessment of relative importance (i.e., a policy decision), of individual infrastructure shortages on the entire collection.

We summarize the multiple-infrastructure defender problem **(MULTI-D)** as follows.

Indices and index sets [~cardinality]

$r \in R$	Infrastructures in a system	(alias *r'*)	[few]
$R(n) \in R$	Infrastructure of node *n*		[few]
$n \in N_r \subseteq N$	Nodes in infrastructure *r*: $N_r \equiv \{n : R(n) = r\}$		[moderate]
$(i, j) \in A_r \subseteq A$	Arcs in infrastructure *r*: $A_r \equiv \begin{Bmatrix} (i, j) \in A : \\ R(i) = R(j) = r \end{Bmatrix}$		[many]

Data [units, if applicable]

h^r	Cost conversion factor given to infrastructure *r* to equalize costs amongst system of infrastructures (cost structure)	[global $/local $]
p^r	System-wide policy weight given to infrastructure r to reflect secondary effects or relative importance of infrastructures (policy decision)	[global $/local $]

Objective (MULTI-D)

$$\min_{Y} f(Y, \bar{X}) \equiv \sum_{r \in R} h^r \left(\sum_{(i,j) \in A_r} \left(c_{ij} + q_{ij} \bar{X}_{ij} \right) Y_{ij} \right) + \sum_{r \in R} p^r \left(\sum_{n \in N_r} \left(ePen_n EXCESS_n + sPen_n SHORT_n \right) \right) \qquad (2.8)$$

$$subject\,to: \quad (2.2) - (2.5)$$

Even though each infrastructure has its own distinct activities, this is essentially a single-commodity minimum cost objective representing the global manager's problem for a collection of infrastructures. The objective function (2.8) now reflects the use of the cost conversion factor (h^r) adjusting the relative network operating costs, as well as the policy weight (p^r) regulating the infrastructure penalties according to policy guidance. Even though this objective has a global perspective, the infrastructures are still

independent, resulting in an objective function that is a sum of the individual weighted infrastructure objective functions. Thus, the model is separable by infrastructure. **MULTI-D** constraints mirror those in **D** (Equations 2.2– 2.5).

The manager for the collection of systems imposes his global view of relative importance through the setting of h^r and p^r. Given these values, the individual infrastructure operator makes locally optimal activity decisions based only on information within his infrastructure, but using the global manager's objective guidance.

2. Attacker Problem (MULTI-AD)

Consider the perspective of an attacker who can allocate his attack resources (Γ) among several independent infrastructures. In the single-infrastructure case, if an attacker possesses the resources for only one attack, he will interdict the activity that maximizes the operator's cost. Assuming all interdictions have the same resource requirements for the attacker, and with multiple infrastructures to consider, the same attacker can now interdict the activity amongst all infrastructures that maximizes the total system operating cost. Therefore, even though the infrastructures are independent, this is a relaxation of **AD** so the attacker can do no worse. Given additional resources, the attacker's influence on the defender's collection of infrastructures is likely to be much greater than if he was only targeting one network.

The defender's problem remains as **MULTI-D**; he is attempting to minimize the weighted costs of his collection of infrastructures. We maintain all attacker-defender assumptions from **AD**; therefore the attacker now has perfect knowledge of the defender's cost conversion factors and policy weights, while the defender will continue to operate the system of infrastructures optimally post-interdiction.

As with the operator's problem, the attacker's objective function for multiple independent infrastructures sums over the weighted infrastructure costs, but does so within the two-stage, attacker-defender **AD** model. To represent an intelligent attacker, the fixed attacks $\bar{X}_{ij}$ are replaced by decision variables X_{ij}. All constraints remain as in **AD** (Equations 2.2–2.5 and 2.7).

C. DIRECT COST-BASED DEPENDENCE

In July 2001, a freight train derailment in the Howard Street Tunnel in Baltimore, Maryland, ignited a chemical fire that burned for days. The incident affected local auto, bus and train transportation, ruptured a main in the water infrastructure, interrupted power to a portion of Baltimore, disrupted East Coast railroad service, and slowed Internet use nationwide due to the destruction of fiber-optic cables in the tunnel serving three of the largest U.S. Internet service providers (DoT, 2002).

This is an example of a geographic (Rinaldi et al., 2001) or co-location (Wallace et al., 2003) dependence, where the disruption in one infrastructure directly affects others, due only to their spatial proximity to a single interdiction. We use costs to represent geographic dependence, but we are not restricted to modeling co-located components. Our formulation allows any cost-based limits on courses of action and activities to be established as needed by the global manager.

1. Defender Problem (DIRECT-D)

To incorporate direct cost-based dependence relationships, we simply expand the interdicted cost in previous models from q_{ij} to q_{ijkl}, so the latter now represents the additional cost to operate on arc $(i, j) \in A$ due to interdiction of arc $(k, l) \in A$. In this manner, a cost-based link may exist between any arc and node in any infrastructure through use of the interdicted-cost, q_{ijkl}. We represent interdicted costs for nodes by splitting the node and assigning the interdicted cost to the internal arc. For simplicity, we also use the term *component* to refer to any node or arc within a network.

Components need not be strictly co-located for an interdiction in one system to affect another. We can depict a local dependence or broader dependence relationships. The flexibility of this method is best displayed through several examples.

To demonstrate local geographic dependence, consider a bridge in a transportation network. In addition to carrying local vehicular traffic, suppose that it also carries a water pipe, electric, telephone and fiber-optic cables over a river. As the loss of

the bridge will also result in the loss of the corresponding pipes and cables, the respective arcs in each of the other four infrastructures will have non-zero interdicted costs.

In addition, we can model dependence on a larger scale (geographic or otherwise), such that flow disruption in an infrastructure affects other networks outside the local area. Consider a severe interruption in local ground transportation as would be caused by the collapse of the Hernando de Soto Bridge in Memphis carrying Interstate 40 across the Mississippi River. This bridge serves as a major east-west artery across the United States and is only one of two crossings over the Mississippi in the Memphis area, serving approximately 45,000–50,000 vehicles a day (MyFox Memphis, 2010). This disruption would require the use of Interstate 55 as an alternate east-west route across the river, already serving 48,000–50,000 vehicles a day. The resulting congestion on Interstate 55 can be captured in an operator's model using an interdiction cost, as the decrease in traffic speeds results in increased operational costs for shipping companies and other infrastructures relying on this stretch of Interstate 40.

Of course, through use of this interdicted cost, we can also represent arc independence; where flow disruption in one infrastructure does *not* affect another in any manner $\left(q_{ijkl} = 0\right)$.

Defining q_{ijkl} requires intimate knowledge of each infrastructure in the global collection in order to accurately portray the cost relationships. Therefore, q_{ijkl} cannot be completely defined by any single infrastructure operator; it must have the oversight of the global manager. In the case of geographic dependence, the manager must first know what components within the collection are co-located, and with the help of individual operators, determine the damage or cost to each infrastructure component if an interdiction occurs in a co-located component. In practice, identifying geographic co-location is a major effort, often requiring the use of geographic information systems (Grubesic & Murray, 2006; Lee et al., 2007; Robert & Marabito, 2010; Bernstein et al., 2011).

We summarize the direct cost-based dependence defender problem **(DIRECT-D)** as follows.

Data [units, if applicable]

q_{ijkl} First-order, per-unit cost of operating on arc $(i, j) \in A$ [cost/unit] induced by interdiction of arc $(k, l) \in A$

Objective (DIRECT-D)

$$\min_{Y} f(Y, \bar{X}) \equiv \sum_{r \in R} h^r \left(\sum_{(i,j) \in A_r} \left(c_{ij} + \sum_{(k,l) \in A} q_{ijkl} \bar{X}_{kl} \right) Y_{ij} \right)$$

$$+ \sum_{r \in R} p^r \left(\sum_{n \in N_r} \left(ePen_n EXCESS_n + sPen_n SHORT_n \right) \right) \qquad (2.9)$$

$$subject\ to: \qquad (2.2) - (2.5)$$

This is a single-commodity, minimum-cost network flow model representing the defender's problem for a collection of infrastructures with direct cost-based dependence relationships. The objective function (Equation 2.9) differs from **MULTI-D** only in the interdicted cost, q_{ijkl}, and represents the minimum cost to operate the system of infrastructures. The constraints remain the same as in prior operator models **D** and **MULTI-D** (Equations 2.2–2.5).

The global manager now must understand cost-based relationships between each infrastructure in his collection. If the only direct dependence relationships are due to co-location, he must be able to define the full set of co-located components for his network operators. The global manager now relies on each individual system operator to define interdicted costs for all nodes and arcs affected by interdicted components $\left(q_{ijkl} \neq 0\right)$. As these interdictable components may or may not be in some operator's own infrastructure, this process can be detailed and cumbersome, but it provides the fidelity necessary to accurately determine collection vulnerabilities due to any cost-based dependence. With knowledge of the known disruptions and interdicted costs, along with the cost conversion factor and policy weight for his infrastructure (provided by the global manager), the infrastructure operator can make his activity decisions without regard to the other infrastructures in the collection.

2. Attacker Problem (DIRECT-AD)

Consider an intelligent attacker wishing to disrupt this system of infrastructures. Given perfect knowledge of the interdicted costs q_{ijkl}, along with penalty costs ($ePen_n$, $sPen_n$) and infrastructure cost conversion factor and policy decisions (h^r, p^r), he can now take advantage of known dependence. The cost-based dependence relationships allow the attacker to impact multiple infrastructures with a single attack, magnifying the potential impact of his attack resources.

Contrary to the attacker's improved situation over **MULTI-AD**, the defender suffers additional vulnerabilities through these direct cost-based dependence relationships. Although the global manager still directs minimization of the collective costs among his infrastructures, the impact of interdictions can now propagate throughout the networks as opposed to interdictions affecting single infrastructures.

DIRECT-AD is identical in formulation to **MULTI-AD**, with the exception of the interdicted costs (q_{ijkl}) and the attack variables (X_{kl}) in the objective function. Constraints remain identical to **AD** (Equations 2.2– 2.5 and 2.7).

D. INDIRECT COMMODITY FLOW DEPENDENCE

The five dependence types (*input, shared, exclusive-or, mutual*, and *co-location*) already referenced from prior literature are defined in terms of relationships between infrastructures, not individual components (Wallace et al., 2003; Lee et al., 2004, 2007). For example, in their definitions, an input dependence refers to an infrastructure requiring one or more services from another infrastructure. In modeling dependence relationships at the component level (node or arc), we make further distinctions between types for completeness. We redefine and model *input*, *shared*, *exclusive-or* and *mutual dependence*, along with introducing two additional types, *substitute* and *complimentary*.

Although some researchers (Rinaldi et al., 2001; Rinaldi, 2004) refer to any situation where one infrastructure depends on commodity output from another infrastructure as a "physical" dependence; we refer to these collectively as an *indirect* dependence. We interpret the cost-based dependence relationships (co-located components for example) from the previous section as *directly* impacting the operator's

objective function. In contrast, the flows from one infrastructure to another only *indirectly* impact the operator's objective function.

1. Derivation of Dependence Type Formulations

Borrowing terminology from Lee et al. (2007), we define all commodity flows between infrastructures as originating at a *parent* node and terminating at a *child* arc. By performing node-splitting, we can model node-node interdependence relationships as needed. We define $S \subseteq N$ as the set of parent nodes providing commodity flow to *supported* infrastructures, and $D \subseteq A$ as the set of child arcs dependent on commodity flow from *supporting* infrastructures. Let $[n,(i,j)] \in G \subseteq S \times D$ be a node-arc pair representing infrastructure dependence. By construction, each child arc $(i,j) \in D$ requires a certain threshold ($threshold_{nij}$) of commodity from its parent node $n \in S$ to support arc operation. Let V_{nij} denote the amount of commodity flow that parent node $n \in S$ provides to arc $(i,j) \in D$. Finally, by definition, if V_{nij} is at least $threshold_{nij}$, then arc (i,j) has the required commodity needed for operation. We use a binary transfer variable T_{nij} to indicate operation of arc $(i,j) \in D$, where $T_{nij} = 1$ allows for the operation of arc $(i,j) \in D$, while $T_{nij} = 0$ indicates commodity flow has not met the threshold required to operate the child arc.

a. Single-Input Dependence

Each indirect, component-level dependence is an *input* dependence according to Wallace et al.'s definition (2003). Therefore, we define a *single-input* dependence at the component level as a single (parent node, child arc) pair $[n,(i,j)] \in G \subseteq S \times D$, with the parent node supporting only one child arc, and the child arc requiring commodity from a single parent node as in Figure 4.

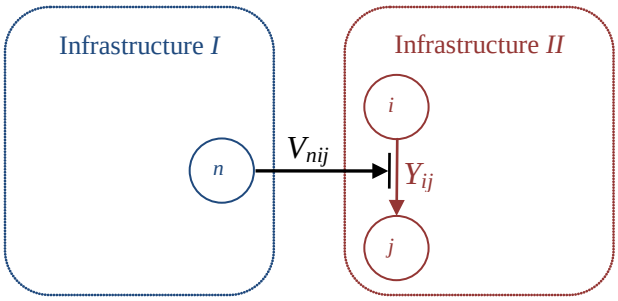

Figure 4. Graphical representation of *single-input* dependence.

Child arc (*i*,*j*) (infrastructure *II*) requires commodity from parent node *n* (infrastructure *I*) to operate. The horizontal arrow indicates dependence between infrastructure components. The black vertical line indicates flow does not enter supported infrastructure as commodity flow, but is necessary for operation of the receiving arc. Arrows indicate direction of flow, both inter- and intra-infrastructure.

As an example of an input dependence, consider an electrically-operated valve in a natural gas transfer system. The valve requires power from an electrical supply to operate and allow transfer of natural gas through the pipeline. Without electrical power, the valve will remain in its default position (closed) and control of the gas is not possible. The electrical system itself receives no direct benefit from supplying power to the valve, but sees the requirement as a demand on the electrical infrastructure.

We capture a *single-input* dependence of a child arc (*i*,*j*) on a parent node *n* with the following additional variable upper bound (VUB) constraints:

$$threshold_{nij} T_{nij} \leq V_{nij} \tag{2.10}$$

$$Y_{ij} \leq u_{ij} T_{nij}. \tag{2.11}$$

Equation (2.10) sets the binary transfer variable $T_{nij} = 1$ only if the commodity flow V_{nij} from parent node *n* is at least the required threshold ($threshold_{nij}$). Equation (2.11) then reduces the upper bound of the child arc flow Y_{ij} to be zero if $T_{nij} = 0$.

The dependence relationship modeled with Equations (2.10) and (2.11) assumes a binary relationship between the child arc (*i*, *j*) flow capacity and commodity flow provided by the parent node *n*. However, there may be physical relationships where

the child arc flow capacity is dependent on the amount of commodity flow provided by the parent node even after activation of the arc $(T_{nij} = 1)$. To model this situation, we require identification of the child arc capacity $\left(\underline{u}_{ij}\right)$ when the minimum threshold flow is provided by the parent node $\left(V_{nij} = threshold_{nij}\right)$, and the commodity flow from the parent node $\left(\overline{V}_{nij}\right)$ that allows for maximum child arc capacity $\left(\overline{u}_{ij}\right)$. Figure 5 provides a comparison of binary and linear input dependence relationships.

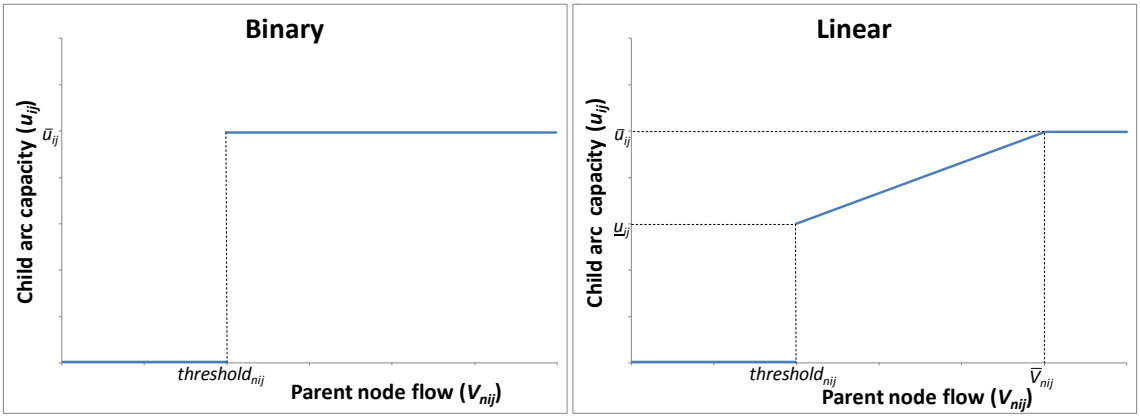

Figure 5. Relationships between commodity flow (V_{nij}) from parent node *n*, and flow capacity (u_{ij}) of child arc (*i*, *j*).

For a binary relationship, if the parent node *n* does not meet the threshold flow needed, the child arc capacity is zero. However, if $V_{nij} \geq threshold_{nij}$, the child arc capacity is $\overline{u}_{ij}$. For a linear relationship, meeting the threshold results in a lower capacity on the child arc $\underline{u}_{ij}$, and further increases in provided flow results in child arc capacity increases, until the maximum arc capacity $\overline{u}_{ij}$ is reached.

We capture the *single-input* dependence with a linear relationship between parent node *n* commodity flow and child arc (*i*, *j*) capacity with Equations (2.10), (2.11) and the additional constraint:

$$Y_{ij} \leq \begin{cases} \underline{u}_{ij} + (\overline{u}_{ij} - \underline{u}_{ij}) \dfrac{(V_{nij} - threshold_{nij})}{(\overline{V}_{nij} - threshold_{nij})} + & \\ \quad (1 - T_{nij}) \left[\dfrac{threshold_{nij} (\overline{u}_{ij} - \underline{u}_{ij})}{(\overline{V}_{nij} - threshold_{nij})} - \underline{u}_{ij} \right] & if \left[\dfrac{threshold_{nij} (\overline{u}_{ij} - \underline{u}_{ij})}{(\overline{V}_{nij} - threshold_{nij})} - \underline{u}_{ij} \right] > 0 \\ \underline{u}_{ij} + (\overline{u}_{ij} - \underline{u}_{ij}) \dfrac{(V_{nij} - threshold_{nij})}{(\overline{V}_{nij} - threshold_{nij})} & otherwise. \end{cases} \quad (2.12)$$

Equation (2.12) sets the child arc capacity for a linear relationship between V_{nij} and u_{ij}. An additional term is required to ensure the capacity remains non-negative for values of V_{nij} less than $threshold_{nij}$.

b. Exclusive-or Dependence

We define an *exclusive-or* dependence as a single parent node $n \in S$ supporting multiple child arcs, but capable of providing the necessary flow to only one child arc $(i, j) \in D$ at a time; see Figure 6. Although this figure depicts child arcs (*i*,*j*) and (*k*,*l*) in separate infrastructures (*II* and *III*, respectively), there may be more than two such child arcs, and some of these children may be in the same infrastructure.

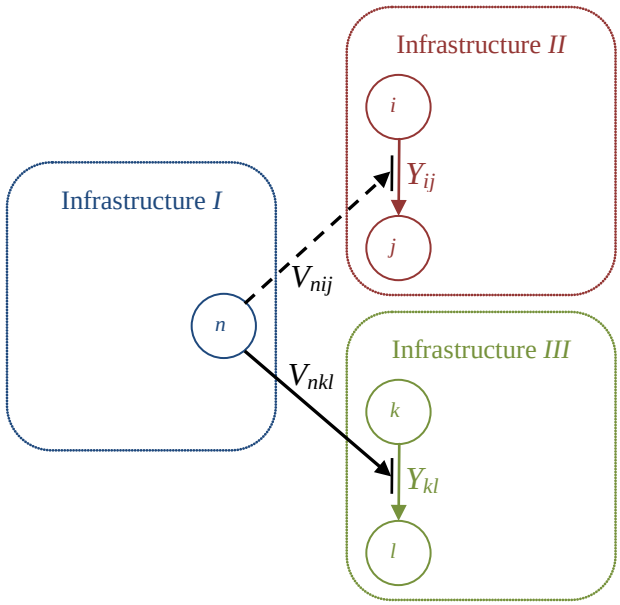

Figure 6. Graphical representation of *exclusive-or* dependence.

Parent node n (infrastructure I) can supply the necessary commodity to either child arc (i,j) (infrastructure II) or child arc (k,l) (infrastructure III) for operation of a single child arc, but not more than one. The dashed black line indicates the infrastructure II dependence that is not being supported, while the solid black line indicates infrastructure III is supported.

Consider as an example of an *exclusive-or* dependence the case where an electrical transfer bus provides power to run a motor or charge a battery. If the bus power is used for the creation of mechanical power, it is not available to charge the battery and vice versa.

We represent an *exclusive-or* dependence between two child arcs (i, j) and (k, l) and a parent node n with the following constraints:

$$threshold_{nij} T_{nij} \le V_{nij} \tag{2.10}$$

$$Y_{ij} \le u_{ij} T_{nij} \tag{2.11}$$

$$threshold_{nkl} T_{nkl} \le V_{nkl} \tag{2.13}$$

$$Y_{kl} \le u_{kl} T_{nkl} \tag{2.14}$$

$$T_{nij} + T_{nkl} \le 1. \tag{2.15}$$

Equations (2.10) and (2.13) set the transfer variable based upon the relationship between commodity flow and required threshold, while Equations (2.11) and (2.14) set the upper bound flow on the respective child arcs based upon the transfer variable value. Equation (2.15) requires a single dependence relationship for *n*. The number of candidate child arcs is not limited to two, and Equation (2.15) can have an arbitrary number of transfer variable terms on its lefthand side.

c. Shared Dependence

We define a *shared* dependence as a single parent node $n \in S$ supplying commodity to multiple child arcs $(i, j) \in D$ and capable of supporting them concurrently, as shown in Figure 7. Let $max\_supportable_n$ represent the maximum number of interdependence links node $n \in S$ can support.

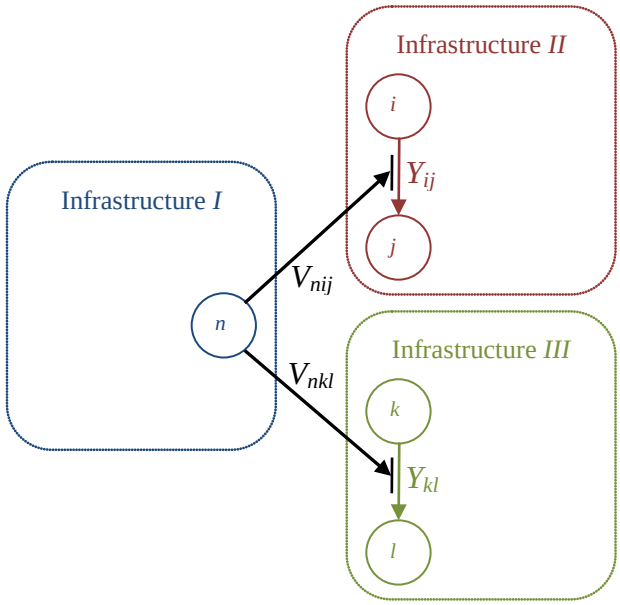

Figure 7. Graphical representation of *shared* dependence.

Arc (*i*,*j*) (infrastructure *II*) and arc (*k*,*l*) (infrastructure *III*) each require commodity from node *n* (infrastructure *I*) to operate.

Consider the case of a homeowner's solar power system. Suppose that he can use the electricity provided by the solar power system to heat his water and meet his electrical needs, while also selling excess generated power to the local electric company.

Any infrastructure component that provides a service to multiple supported components simultaneously serves as an example of a *shared* dependence.

We represent a *shared* dependence between two child arcs (i, j) and (k, l) and a parent node n in a similar manner to exclusive-or, maintaining Equations (2.10) through (2.14) to set the transfer variables and control the upper bound on dependent commodity flows. We replace Equation (2.15) with (2.16) to restrict the number of dependence relationships $n \in S$ can support.

$$\sum_{(i,j):(n,(i,j))\in G} T_{nij} \leq max\_supportable_n \tag{2.16}$$

Equation (2.16) can have an arbitrary number of transfer variable terms.

d. Substitute Dependence

We define *substitute* dependence as a single child arc $(k,l) \in D$ requiring commodity from at least one of several parent nodes $n \in S$ to operate, as shown in Figure 8.

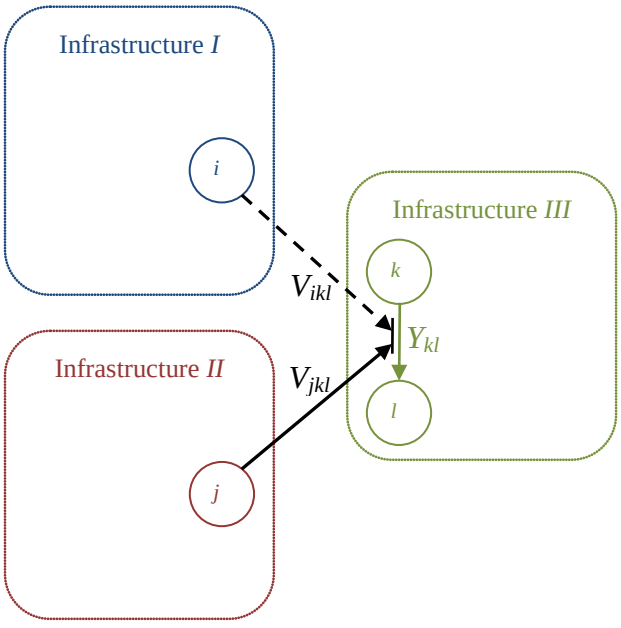

Figure 8. Graphical representation of *substitute* dependence.

Either node i (infrastructure I) or node j (infrastructure II) must supply arc (k,l) (infrastructure III) with needed commodity for operation.

As an example of *substitute* dependence, consider a water system's pumping station that can be operated either by electricity from the grid, or from a gasoline-powered backup generator.

We represent *substitute* dependence of a child arc (*k*,*l*) on two substitutable parent nodes *i* and *j* with the following additional constraints:

$$threshold_{ikl} T_{ikl} \le V_{ikl} \tag{2.17}$$

$$threshold_{jkl} T_{jkl} \le V_{jkl} \tag{2.18}$$

$$Y_{kl} \le u_{kl}(T_{ikl} + T_{jkl}). \tag{2.19}$$

Equations (2.17) and (2.18) set the respective transfer variables independently based upon available flow from each parent node. The commodity flow upper bound on the child arc is set with constraint (2.19) by summing the transfer variables and multiplying by the arc capacity. The normal arc capacity constraint (Equation 2.3) still applies, such that the arc flow Y_{kl} will not exceed the arc capacity, even if each parent node supplies the necessary commodity.

e. Complimentary Dependence

We define a *complimentary* dependence as a single child arc $(k,l) \in D$ requiring commodity from more than one parent node $n \in S$ to operate, as shown in Figure 9. Let $min\_required_{kl}$ represent the minimum number of dependence links that arc $(k,l) \in D$ requires. We also require another binary variable W_{kl}, to indicate whether all required threshold commodities needed for operation are provided. Consider the case where $T_{ikl} = 1$ and $T_{jkl} = 1$, indicating nodes *i* and *j* are meeting their respective required thresholds ($threshold_{ikl}$ and $threshold_{jkl}$). If both commodities are required for operation of arc $(k,l) \in D$, then $W_{kl} = 1$ if and only if both $T_{ikl} = 1$ and $T_{jkl} = 1$.

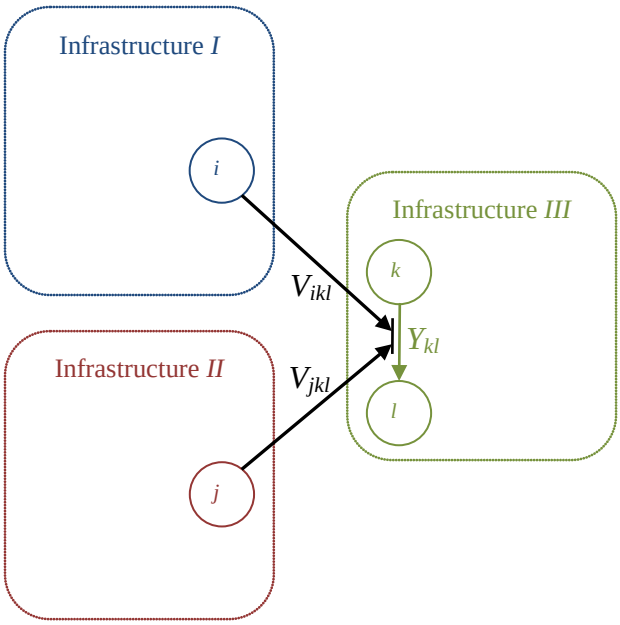

Figure 9. Graphical representation of *complimentary* dependence.

Both node *i* (infrastructure *I*) and node *j* (infrastructure *II*) must supply arc (*k*,*l*) (infrastructure *III*) with needed commodity for its operation.

As an example of a *complimentary* dependence, consider a water system pumping station that needs both electricity and water to operate. We represent a *complimentary* dependence of child arc (*k*, *l*) on two parent nodes *i* and *j* with Equations (2.17) and (2.18) along with the following additional constraints:

$$\sum_{n:(n,(k,l))\in G} T_{nkl} \geq min\_required_{kl} W_{kl} \tag{2.20}$$

$$Y_{kl} \leq u_{kl} W_{kl}. \tag{2.21}$$

Equations (2.17) and (2.18) set the individual transfer variables (T_{ikl} and T_{jkl}, while constraint (2.20) sets the child arc transfer variable (W_{kl}) based upon required dependence relationships. We set the arc flow upper bound with Equation (2.21) in concert with normal arc capacity constraint (2.3).

f. Mutual Dependence

A *mutual* dependence at the infrastructure level rarely translates to a mutual dependence at the component level. Because we model dependence between

parent nodes and child arcs, mutual dependence can only arise between node pairs in separate infrastructures, where each node relies on commodity flow from the other. We model this through use of node-splitting as shown in Figure 10.

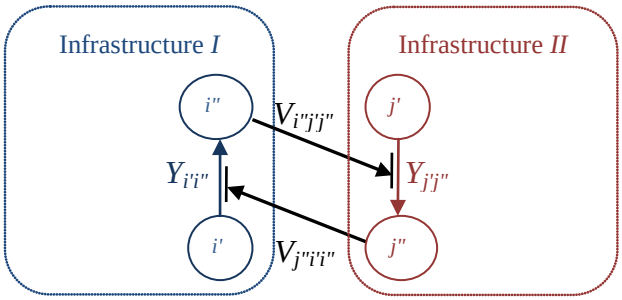

Figure 10. Graphical representation of *mutual* dependence.

Node *i* (infrastructure *I*) and node *j* (infrastructure *II*) each depend on commodity from the other infrastructure to operate. To model, we perform node-splitting and represent this as node-arc dependence.

Consider an electric power generator directly supplying a pump or compressor in a natural gas network. If the pump is also the source of natural gas supply for the generator, a mutual dependence exists.

We represent a *mutual* dependence of child arcs (*i'*, *i"*) and (*j'*, *j"*) on parent nodes *j"* and *i"* respectively using the following additional constraints:

$$threshold_{i''j'j''} T_{i''j'j''} \le V_{i''j'j''} \quad (2.22)$$

$$threshold_{j''i'i''} T_{j''i'i''} \le V_{j''i'i''} \quad (2.23)$$

$$Y_{j'j''} \le u_{j'j''} T_{i''j'j''} \quad (2.24)$$

$$Y_{i'i''} \le u_{i'i''} T_{j''i'i''}. \quad (2.25)$$

Equations (2.22) and (2.23) set the respective transfer variables to allow flow if thresholds are met, while constraints (2.24) and (2.25) set the child arc flow upper bounds.

2. Defender Problem (INDIRECT-D)

With indirect dependence relationships present between infrastructures, the global manager now has a more complex system to consider. We assume each infrastructure operator knows both his parent node and child arc dependence relationships. The operators report the status of their parent nodes to the global manager and he in turn passes the status of dependence links to supported infrastructure operators so that they may operate their individual networks efficiently.

Through the independent formulations of our six dependence types, we summarize all notation required for the formulation of the complete dependence defender problem (**INDIRECT_D**) as follows.

Indices and index sets

$S \subseteq N$	Nodes supplying commodity flow to another infrastructure	[few]
$D \subseteq A$	Arcs that depend on commodity flow from another infrastructure	[few]
$\left[n,(i,j)\right] \in G \subseteq S \times D$	Node-arc pair representing infrastructure interdependence	[few]

Data [units, if applicable]

$threshold_{nij}$	Threshold of input commodity needed at arc $(i,j) \in A$ from node $n \in S$	[commodity unit]
$max\_supportable_n$	Number of dependence links that node $n \in S$ can support	[cardinality]
$min\_required_{kl}$	Number of dependence links that arc $(i,j) \in D$ requires for operation	[cardinality]

Decision variables [units, if applicable]

T_{nij}	Variable indicating whether node $n \in S$ is providing the threshold commodity needed for operation of arc $(i, j) \in D$	[binary]
W_{ij}	Variable indicating whether arc $(i, j) \in D$ is receiving all required threshold commodities needed for operation	[binary]
V_{nij}	Flow variable representing commodity flow from node $n \in S$ to support operation of arc $(i, j) \in D$	[commodity unit]

Formulation of Defender Problem (INDIRECT-D)

$$\min_{\substack{Y \\ T,W,V}} f(Y,\bar{X}) \equiv \sum_{r\in R} h^r \left(\sum_{(i,j)\in A_r} \left(c_{ij} + \sum_{(k,l)\in A} q_{ijkl}\bar{X}_{kl} \right) Y_{ij} \right) \tag{2.9}$$

$$+ \sum_{r\in R} p^r \left(\sum_{n\in N_r} \left(ePen_n EXCESS_n + sPen_n SHORT_n \right) \right)$$

subject to :

$$\sum_{j:(n,j)\in A} Y_{nj} - \sum_{i:(i,n)\in A} Y_{in} + \sum_{(i,j):(n,(i,j))\in G} V_{nij}$$

$$+ EXCESS_n - SHORT_n = b_n \qquad \forall n \in N \tag{2.26}$$

$$threshold_{nij} T_{nij} \le V_{nij} \qquad \forall (n,(i,j)) \in G \tag{2.27}$$

$$Y_{ij} \le u_{ij} W_{ij} \qquad \forall (i,j) \in D \tag{2.28}$$

$$SHORT_n \le -b_n + \sum_{(i,j):(n,(i,j))\in G} V_{nij} \qquad \forall n \in N \tag{2.29}$$

$$\sum_{(i,j):(n,(i,j))\in G} T_{nij} \le max\_supportable_n \qquad \forall n \in S \tag{2.30}$$

$$\sum_{n:(n,(i,j))\in G} T_{nij} \ge min\_required_{ij} W_{ij} \qquad \forall (i,j) \in D \tag{2.31}$$

$$0 \le Y_{ij} \le u_{ij} \qquad \forall (i,j) \in A \tag{2.3}$$

$$EXCESS_n \ge 0,\ SHORT_n \ge 0 \qquad \forall n \in N \tag{2.4}$$

$$Y \in \Psi \tag{2.5}$$

This is a single-commodity minimum cost network flow model representing the defender's problem for a collection of infrastructures with direct cost-based and indirect flow-based dependence relationships. The objective function for **INDIRECT-D** is identical to **DIRECT-D**, as are Equations (2.3), (2.4) and (2.5). The balance of flow equation (2.26) adds the dependence flow V_{nij} for parent nodes. Equation (2.27) sets the dependence threshold variable T_{nij} for all system dependence relationships, and child arc flow upper bounds are set with constraint (2.28). Demand node shortages are adjusted to allow for commodity flow between infrastructures at parent nodes (2.29). Constraint (2.30) restricts the number of child arcs a parent node can support, while equation (2.31)

sets the child arc transfer variable based upon the required number of dependence links required for operation. While Equation (2.9) explicitly shows the minimization is with regard to variables *Y, T, W* and *V*, we will subsequently write "$\min_{Y}$" for brevity with the understanding that we are also minimizing with respect to the *T, W,* and *V* variables (and *EXCESS* and *SHORT*) as well.

3. Attacker Problem (INDIRECT-AD)

We now consider an intelligent attacker solving for the worst case global interdiction possible with his attack resources. He can take advantage of direct cost-based and indirect commodity dependence relationships to maximize the global manager's minimum cost. We assume the attacker's perfect knowledge extends to all dependence relationships, both direct and indirect.

We summarize the complete dependence attacker problem (**INDIRECT-AD**) as follows.

Formulation of Attacker Problem (INDIRECT-AD)

$$\max_{X} \min_{Y} f(Y,X) \equiv \sum_{r\in R} h^r \left(\sum_{(i,j)\in A_r} \left(c_{ij} + \sum_{(k,l)\in A} q_{ijkl} X_{kl} \right) Y_{ij} \right) \tag{2.32}$$

$$+ \sum_{r\in R} p^r \left(\sum_{n\in N_r} \left(ePen_n EXCESS_n + sPen_n SHORT_n \right) \right)$$

subject to:

$$\sum_{j:(n,j)\in A} Y_{nj} - \sum_{i:(i,n)\in A} Y_{in} + \sum_{(i,j):(n,(i,j))\in G} V_{nij}$$

$$+ EXCESS_n - SHORT_n = b_n \qquad \forall n \in N \tag{2.26}$$

$$threshold_{nij} T_{nij} \le V_{nij} \qquad \forall (n,(i,j)) \in G \tag{2.27}$$

$$Y_{ij} \le u_{ij} W_{ij} \qquad \forall (i,j) \in D \tag{2.28}$$

$$SHORT_n \le -b_n + \sum_{(i,j):(n,(i,j))\in G} V_{nij} \qquad \forall n \in N \tag{2.29}$$

$$\sum_{(i,j):(n,(i,j))\in G} T_{nij} \le max\_supportable_n \qquad \forall n \in S \tag{2.30}$$

$$\sum_{n:(n,(i,j))\in G} T_{nij} \ge min\_required_{ij} W_{ij} \qquad \forall (i,j) \in D \tag{2.31}$$

$$0 \le Y_{ij} \le u_{ij} \qquad \forall (i,j) \in A \tag{2.3}$$

$$EXCESS_n \ge 0,\ SHORT_n \ge 0 \qquad \forall n \in N \tag{2.4}$$

$$Y \in \Psi \tag{2.5}$$

$$X \in \Gamma \tag{2.6}$$

INDIRECT-AD represents the attacker's problem for a single-commodity minimum cost network flow model of a collection of fully interdependent infrastructures. The objective function (2.32) expresses with respect to *X* the operating cost achievable with respect to commodity flow *Y*. Constraints (2.26) thru (2.31) are identical to **INDIRECT-D**, Equation (2.3), (2.4) and (2.5) with all prior models, and Equation (2.6) is consistent with prior **AD** formulations.

E. SOLVING INDIRECT-AD WITH DECOMPOSITION

We solve **INDIRECT-AD** with Benders decomposition as follows. Our subproblem is simply the defender's problem, **INDIRECT-D**, for a fixed set of interdictions $\bar{X}$. Each such subproblem yields an optimal set of defender flows Y^* over the collection of infrastructure systems. At each iteration *m* of our algorithm, we record this optimal set of flows as $\bar{Y}^m$ (with associated values $\overline{EXCESS}_n^m$ and $\overline{SHORT}_n^m$). Each of these solutions yields a bound on the total system cost *Z*, which the attacker can force the defender to pay:

$$Z \le \sum_{r \in R} h^r \left(\sum_{(i,j) \in A_r} \left(c_{ij} + \sum_{(k,l) \in A} q_{ijkl} X_{kl} \right) \bar{Y}_{ij}^m \right)$$

$$+ \sum_{r \in R} p^r \left(\sum_{n \in N_r} \left(ePen_n \overline{EXCESS}_n^m + sPen_n \overline{SHORT}_n^m \right) \right) \qquad \forall m \in M. \qquad (3.33)$$

The collection of these bounds, in addition to any constraints on the interdictions themselves, yields the attacker's master problem **MASTER-AD**.

We summarize the Benders formulation of the master problem as follows.

Indices and index sets

$m \subseteq M$	Decomposition iteration counter	[few]

Data [units, if applicable]

$\bar{Y}_{ij}^m$	Optimal operator's flow plan (solved in subproblem) for iteration *m*	[commodity unit]
$\overline{EXCESS}_n^m$	Excess commodity at node *n* for optimal operator plan in iteration *m*	[commodity unit]
$\overline{SHORT}_n^m$	Commodity shortage at node *n* for optimal operator plan in iteration *m*	[commodity unit]

Formulation of Attacker Problem (MASTER-AD)

$$Z_{\max}(\overline{Y}) = \max_{X} Z \tag{3.34}$$

subject to:

$$Z \le \sum_{r \in R} h^r \left(\sum_{(i,j) \in A_r} \left(c_{ij} + \sum_{(k,l) \in A} q_{ijkl} X_{kl} \right) \overline{Y}_{ij}^{m} \right)$$

$$+ \sum_{r \in R} p^r \left(\sum_{n \in N_r} \left(ePen_n \, \overline{EXCESS}_n^m + sPen_n \, \overline{SHORT}_n^m \right) \right) \qquad \forall m \in M \tag{3.33}$$

The objective function (3.34) evaluates the attacker's plan responding to the optimal operator's commodity flow plan $\left(\overline{Y}\right)$. Each subproblem solution (**INDIRECT-D**) provides an additional constraint (3.33) for each iteration *m*.

The complete decomposition algorithm follows:

Algorithm MASTER-AD

Input: infrastructure data, attacker resources, optimality tolerance $\varepsilon \geq 0$.

Output: ε-optimal attack plan X^*, responding operator plan Y^*.

1. Initialize best lower bound $Z_{LB} \leftarrow -\infty$, best upper bound $Z_{UB} \leftarrow +\infty$, define the incumbent, null attack plan $\bar{X}^1 \leftarrow 0$ as the best found so far, and set iteration counter $M \leftarrow 1$.
2. **Subproblem**: use attacker plan $\bar{X}^M$ to solve **INDIRECT-D** and determine the optimal operator's responding activity plan $\bar{Y}^M$. The bound on the associated objective is $Z_{\min}\left(\bar{X}^M\right)$.
3. If $M = 1$ and $\bar{X}^1 \notin X$ (i.e., not admissible), go to Step 6 (**Master Problem**).
4. If $Z_{LB} < Z_{\min}\left(\bar{X}^M\right)$ set $Z_{LB} \leftarrow Z_{\min}\left(\bar{X}^M\right)$ and record improved incumbent attack plan $X^* \leftarrow \bar{X}^M$, and responding operator plan $Y^* \leftarrow \bar{Y}^M$.
5. If $Z_{UB} - Z_{LB} \leq \varepsilon$ go to **END**.
6. **Master Problem**: use operator plans $\bar{Y}^M$ to solve MASTER-AD and determine an optimal attacker plan $\bar{X}^{M+1}$. The bound on the associated objective is $Z_{\max}\left(\bar{Y}\right)$.
7. If $Z_{UB} > Z_{\max}\left(\bar{Y}\right)$ set $Z_{UB} \leftarrow Z_{\max}\left(\bar{Y}\right)$.
8. If $Z_{UB} - Z_{LB} \leq \varepsilon$ go to **END**.
9. Set $M \leftarrow M + 1$ and go to Step 2 (**Subproblem**).
10. **END**: X^* is the ε-optimal attack plan, and Y^* is the responding operator plan.

III. MODEL DEMONSTRATION

With MASTER-AD fully defined, we clarify necessary concepts before introducing a small collection of infrastructures to demonstrate the model's use.

We make use of three separate decision-makers in these demonstrations. An *operator* is the owner of an individual infrastructure within the collection, while the *manager* is the supervisor of the entire collection. The *operator* maintains a local perspective of his infrastructure, while the *manager* has a global interest in the collection as a whole. In keeping with the attacker-defender construct, the global *manager* is also referred to as the *defender* (attempting to minimize collection operation costs), and the *attacker* is the adversary attempting to maximize these same costs.

In Chapter II, we introduced the concept of direct cost-based dependence through the use of q_{ijkl}, and we discussed the importance of carefully defining these costs to achieve desired model results that mirror reality. For example, if an operator does not have the ability to send commodity flow past a failed electrical valve, our optimal model solution must not suggest flow across this valve if it is interdicted. Conversely, if the operator has a backup battery for the electric valve, our model must allow for flow across the interdicted valve. By setting q_{ijkl} carefully, we ensure our model reflects the desired behavior. The following scenarios make use of each of these situations.

Lastly, we define *flow disruption* as the case where there is zero commodity flow across a component, infrastructure, or collection (as indicated) due to interdiction. Flow disruption across an infrastructure indicates commodity is no longer shipped from any supply nodes within that infrastructure due to attack, while flow disruption across the collection indicates the attacker has succeeded in stopping all flow within the collection of infrastructures.

A. MULTIPLE INDEPENDENT INFRASTRUCTURES

We first consider three independent infrastructures (denoted here as *r1, r2, r3*), each consisting of three nodes and three arcs, and each managed by an individual operator who attempts to satisfy supply and demand at minimum cost. By construction,

during normal operation with no interdictions, the minimum-cost path for each infrastructure is a direct path between the supply and demand nodes, although one alternate (more expensive) path exists through a transshipment node as shown in Figure 11. For example, in infrastructure *r2*, there is a supply of 10 units at node *r2n1* and a demand of 10 units at node *r2n3*. The low-cost path (at $8 per unit of flow) is directly from *r2n1* to *r2n3*. There is a secondary path passing through node *r2n2*, but it costs a total of $10 per unit flow.

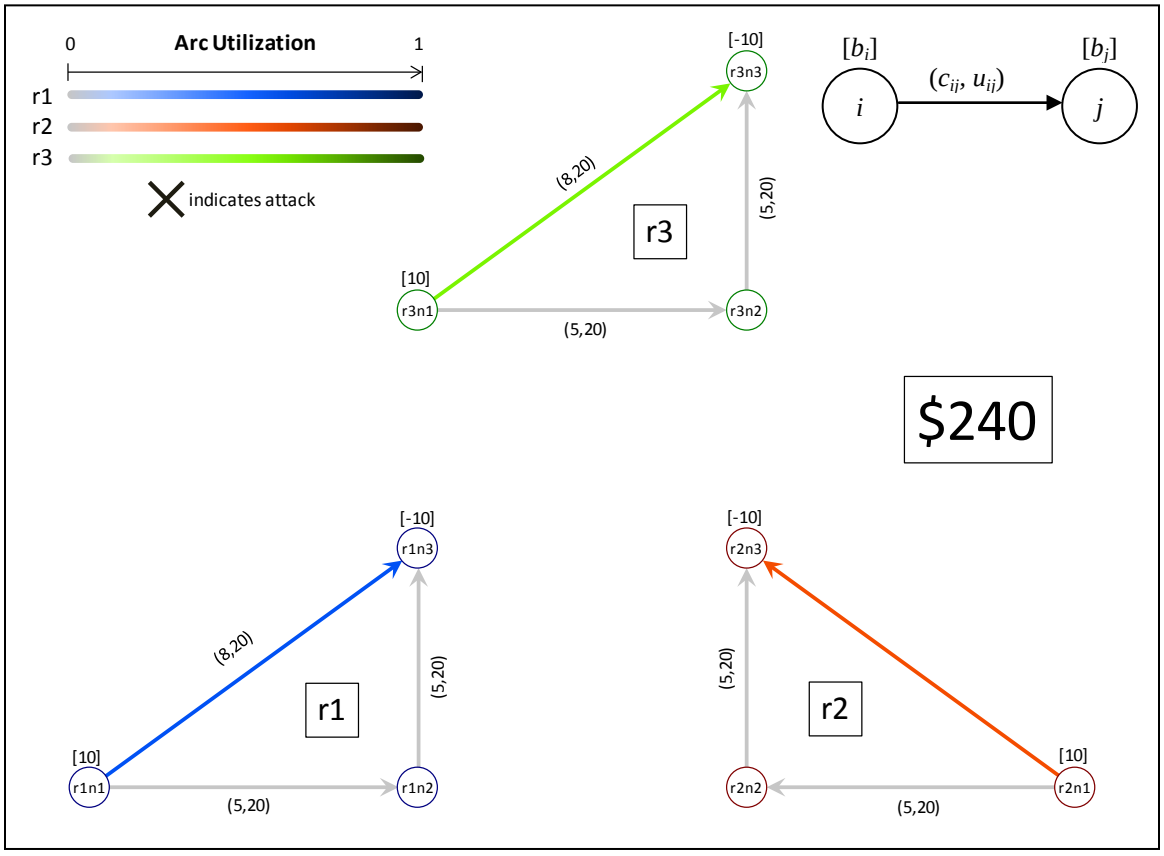

Figure 11. Multiple *independent* infrastructures during normal operation.

Notation is shown in upper right. Colored arcs indicate commodity flow, grey arcs have no flow. Flow within each infrastructure is shown in a separate color with intensity representing arc capacity utilization as indicated in the legend. During normal operation of the collection with no interdictions, commodity flow is direct from the supply to demand nodes in all three infrastructures, with a minimum-cost objective of $240.

1. Model Input

Consider the case where all infrastructure costs are measured in dollars $\left(h^r = 1, \forall r \in R\right)$ and the global manager assesses the secondary effects and relative policy importance of each infrastructure to be the same $\left(p^r = 1, \forall r \in R\right)$. In addition, assume that each infrastructure operator, recognizing the lack of robustness of a single supply and demand system, has hardened his supply and demand nodes, making them invulnerable to attack. However, suppose that all arcs and transshipment nodes across the collection of infrastructures are vulnerable to attack. Assume that each infrastructure has sufficient storage capability, so no penalties are charged for excess commodity availability, but penalties for shortages are assessed equally at \$15 per unit of commodity. In addition, assume that the per-unit cost to operate across interdicted arcs is \$10 globally, while the cost to operate an interdicted node is set higher, at \$25. For our example, these costs are high enough to prevent shipment of commodity across interdicted components. Also, no direct dependence relationships exist; interdicted arcs and nodes do not have a direct cost effect on any other arcs or nodes $\left[q_{ijkl} = 0, \forall (i, j) \neq (k, l)\right]$. We summarize the model input in Table 1.

System Data

r	h^r	p^r
r1	1	1
r2	1	1
r3	1	1

Node Data

n	Vuln	b_n	$ePen_n$	$sPen_n$
r1n1		10	0	
r1n2	1			
r1n3		-10		15
r2n1		10	0	
r2n2	1			
r2n3		-10		15
r3n1		10	0	
r3n2	1			
r3n3		-10		15

Arc Data

i	j	Vuln	c_{ij}	u_{ij}
r1n1	r1n2	1	5	20
r1n1	r1n3	1	8	20
r1n2	r1n3	1	5	20
r2n1	r2n2	1	5	20
r2n1	r2n3	1	8	20
r2n2	r2n3	1	5	20
r3n1	r3n2	1	5	20
r3n1	r3n3	1	8	20
r3n2	r3n3	1	5	20

Interdiction Data (q_{ijkl})

i	j	k	l	q_{ijkl}
r1n2	r1n2	r1n2	r1n2	25
r2n2	r2n2	r2n2	r2n2	25
r3n2	r3n2	r3n2	r3n2	25
r1n1	r1n2	r1n1	r1n2	10
r1n1	r1n3	r1n1	r1n3	10
r1n2	r1n3	r1n2	r1n3	10
r2n1	r2n2	r2n1	r2n2	10
r2n1	r2n3	r2n1	r2n3	10
r2n2	r2n3	r2n2	r2n3	10
r3n1	r3n2	r3n1	r3n2	10
r3n1	r3n3	r3n1	r3n3	10
r3n2	r3n3	r3n2	r3n3	10

Table 1. Model Input – multiple *independent* infrastructures.

A "1" in the "Vuln" columns (*Node Data* and *Arc Data* sections) indicates the component is vulnerable to attack. Each vulnerable component has an interdicted cost shown in the *Interdiction Data* section. Blank entries are zero.

2. Initial Results

Based on these model parameters and costs, if the attacker can afford a single attack $\left(\text{i.e., } \Gamma = \left\{ X : \sum_{(i,j)\in A} X_{ij} \le 1 \right\}\right)$, an optimal attack plan is to disrupt the low-cost direct arc between the supply and demand nodes in an infrastructure. Because the three infrastructures are identical, the attacker can select an infrastructure arbitrarily. If the attacker can afford two attacks, the optimal attack plan is to disrupt both the arc between supply and demand nodes and the single transshipment node (the cost of operation on an interdicted node is greater than for an interdicted arc) within an infrastructure. Because it costs less to suffer a shortage in an infrastructure than ship commodity across an interdicted arc or node, the worst two-component attack is one that results in total flow disruption across a single infrastructure, as shown in Figure 12. By extension, the worst six-component attack is one that results in total flow disruption across the entire collection of infrastructures (not shown).

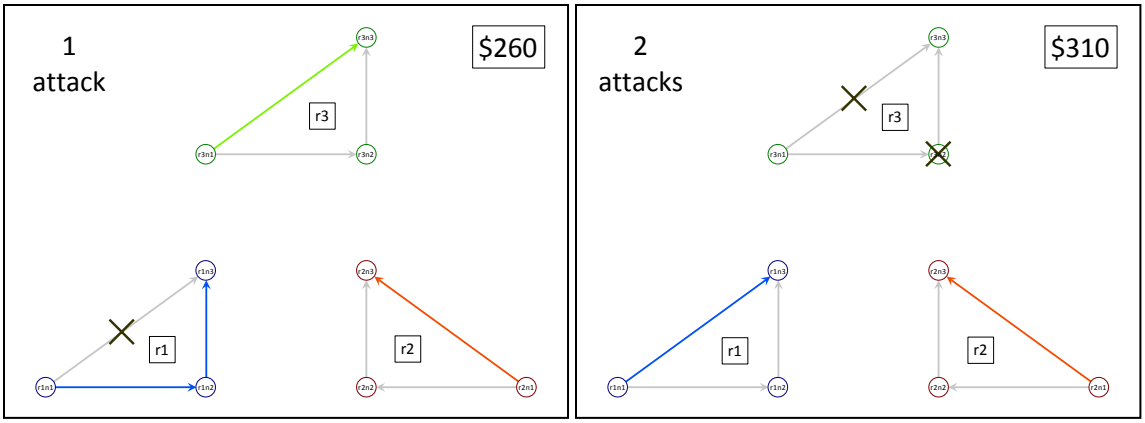

Figure 12. Model Results for multiple *independent* infrastructures.

Each panel shows the location of the optimal attacks and subsequent infrastructure operation. The attack resources available are listed in the upper left corner and the resulting objective function value is boxed in the upper right corner. A single attack resource allows an intelligent attacker to attack the low-cost path in any of the three infrastructures, requiring selection of the alternate, higher-cost path by the operator to flow commodity. When attacking the alternate path, the attacker chooses to interdict the transshipment node because the cost for the defender to operate an interdicted node ($25) is higher than the cost to operate on an interdicted arc ($10). Two attacks result in total flow disruption of a single infrastructure. With all infrastructures identical and all cost conversion factors and policy weights equal to one, selection of the infrastructures for attack is arbitrary.

The global manager's operating cost for this collection of infrastructures increases as the attacker resources increase, until all flow across the global collection is disrupted when the attacker can afford six attacks (Figure 13).

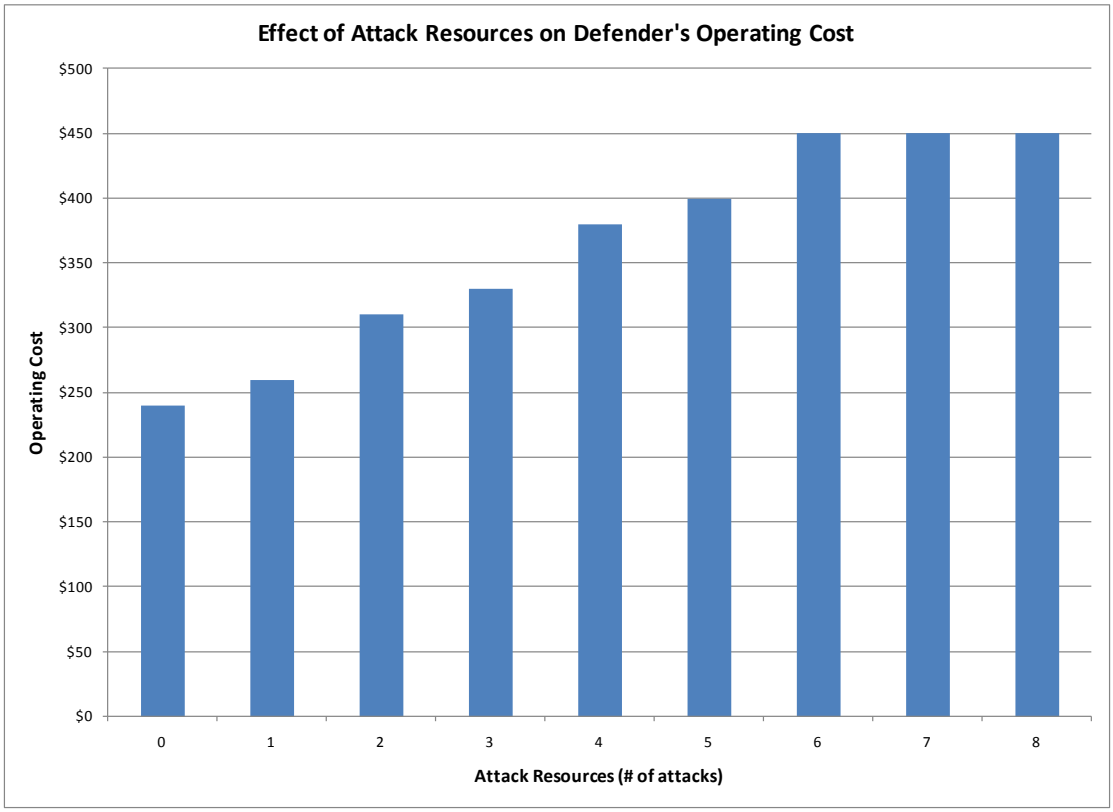

Figure 13. Effect of attack resources on a defender's operating cost for a collection of *independent* infrastructures.

All commodity flow across the collection of infrastructures is disrupted with six attacks, so no additional cost is incurred by the operator for further increases in attack resources. Even with additional resources, the attacker does not benefit from more than six attacks.

3. Effects of Cost Conversion Factors and Policy Weights

Using the previous example as a "base case," we demonstrate the use and effect of the cost conversion factors and policy weights. Consider the case where the objective functions for the three individual infrastructures are measured in slightly different cost units, requiring cost conversions using h^r. In addition, assume that unmet demand between infrastructures results in secondary impacts on society, which are reflected in the policy weights. We display modified model input in Table 2.

System Data

r	h^r	p^r
r1	1	**1.4**
r2	**1.2**	**1.2**
r3	**1.4**	1

Table 2. Modified Model Input – updated cost conversion factors (h^r) and policy weights (p^r) for multiple *independent* infrastructures.

When considering the effects of h^r and p^r in isolation from each other, a disruption in the operation of *r3* is most costly $\left(h^{r3} > h^{r2} > h^{r1}\right)$ while commodity shortages in *r1* are most costly $\left(p^{r1} > p^{r2} > p^{r3}\right)$. However, when considering both cost conversions and policy weights together, the results are not completely intuitive, as shown in Figure 14.

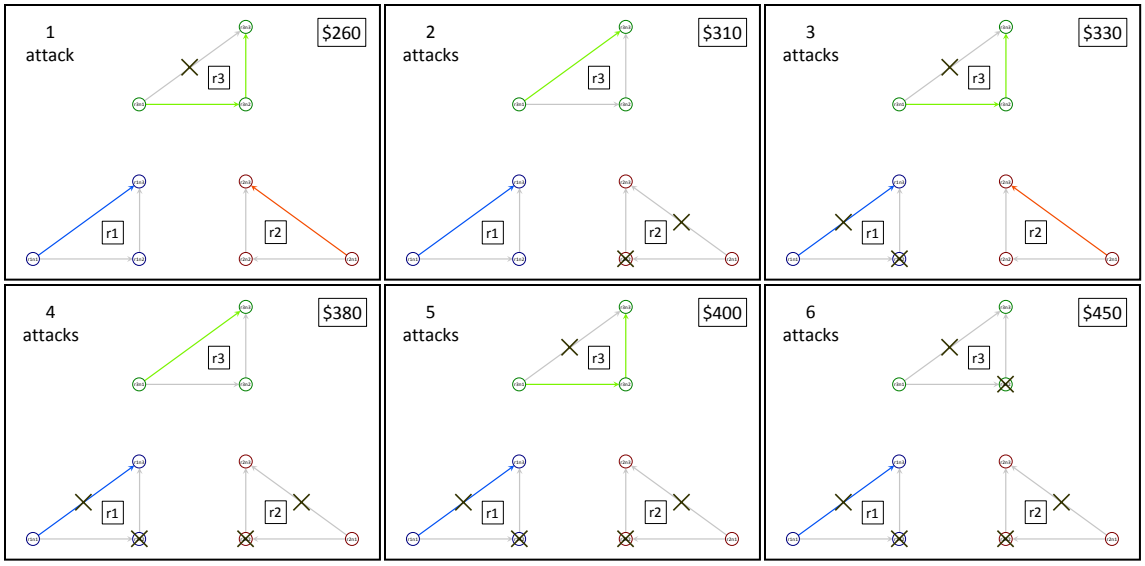

Figure 14. Model Results for updated cost conversions and policy weights for multiple *independent* infrastructures.

Infrastructure *r3* has highest cost factor, resulting in the minimum-cost path in *r3* as the optimal single-attack. Two attacks result in flow disruption in *r2*, even though *r1* has the highest policy weight of all infrastructures. The three-attack results illustrate the cost conversion factor and policy weight in infrastructure *r1* result in optimally shipping commodity across an interdicted arc for an additional cost as opposed to suffering demand shortages. Therefore, the attacker can no longer disrupt flow across the entire collection (only *r2* and *r3*), and he gains no additional benefit with more than six attacks.

This test case illustrates two important points. First, p^{r1} is sufficiently large that the optimal plan for a global manager is to send commodity flow across an interdicted arc (at an additional cost) rather than taking a shortage penalty in *r1*. Therefore, if an attacker can afford two attacks, his optimal plan is to disrupt flow in *r2* rather than *r1*, as *r2* has the second highest policy weight of the collection. This demonstrates the attacker-defender premise that the optimal attack plan assumes the subsequent operation by the global manager to be optimal.

Secondly, in the absence of a global manager, the *r1* operator makes myopic decisions that are locally optimal, but globally suboptimal. If given no guidance from the global system manager, the operator would not ship across an interdicted arc or node. However, due to the cost conversion factors and policy weights, the global manager ships

commodity across an interdicted arc in *r1* for optimal operation of the collection. This case reinforces our assertion that a global manager, as opposed to individual infrastructure operators, is the only one who can set policy weights and drive decisions required for the overall good of the infrastructures.

B. COLLECTION OF INFRASTRUCTURES WITH CO-LOCATED COMPONENT

We now consider direct dependence relationships and their impact on attacker and defender decisions when analyzing a collection of infrastructures. As the Howard Street Tunnel accident in Baltimore illustrated, the co-location of components can have severe consequences when disruptions occur. Adding a co-location dependence relationship to a collection of infrastructures can introduce a new way for an attacker to indirectly influence the system.

We return to the base case defined previously with equal cost and policy weights (all set to one). However, we now additionally consider a single direct dependence, defined by two arcs that are geographically co-located, so an attack on either arc is an interdiction of both (Figure 15). This dependence results in the addition of two interdiction costs as additional model input (Table 3).

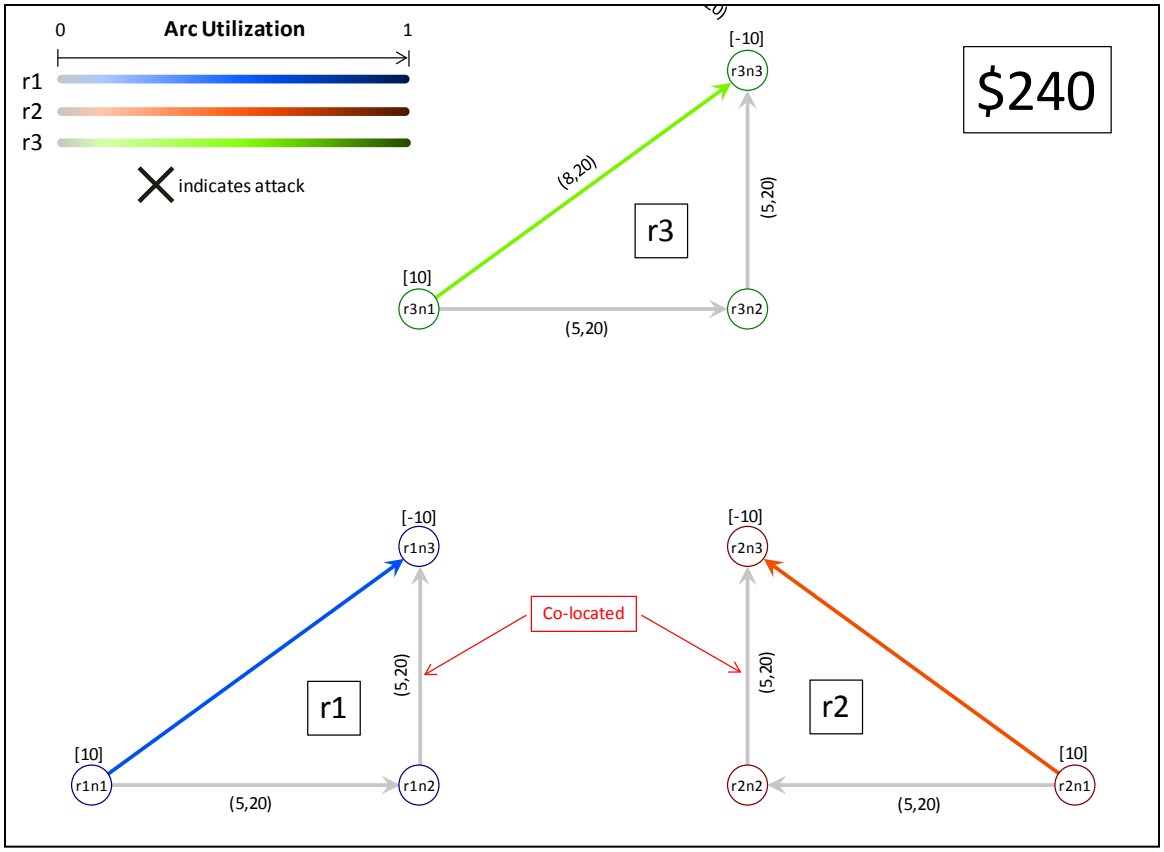

Figure 15. *Direct* Dependence. Three identical infrastructures shown during normal operation.

Arcs (*r1n2*, *r1n3*) and (*r2n2*, *r2n3*) are geographically co-located; therefore, an attack on either results in an interdiction cost on both.

System Data

r	h^r	p^r
r1	1	1
r2	1	1
r3	1	1

Interdiction Data

i	j	k	l	q_{ijkl}
r1n2	r1n3	r2n2	r2n3	10
r2n2	r2n3	r1n2	r1n3	10

Table 3. Model Input – *direct* dependence.

System Data now reflects cost conversion factors and policy weights that are all equal to one, while *Interdiction Data* shows additional interdiction costs (q_{ijkl}) for co-located components. For example, an interdiction of (*r1n2*, *r1n3*) now results in an interdiction cost of $10 per unit of commodity flow on (*r2n2*, *r2n3*) and vice versa. All other input data (node, arc and interdiction) remains unchanged from the base case shown in Table 1.

The addition of the direct dependence to the model has the effect of giving the attacker an additional resource, because with a single attack on either (*r1n2*, *r1n3*) or (*r2n2*, *r2n3*) he gets the benefit of interdicting both arcs. In the base case, an attacker could disrupt all flow across any infrastructure with two attacks and disrupt flow across the entire collection with six attacks. However, due to the direct dependence, the attacker can now disrupt all flow in *r1* and *r2* with only three attacks, and disrupt flow across the entire collection of infrastructures with five attacks (Figure 16).

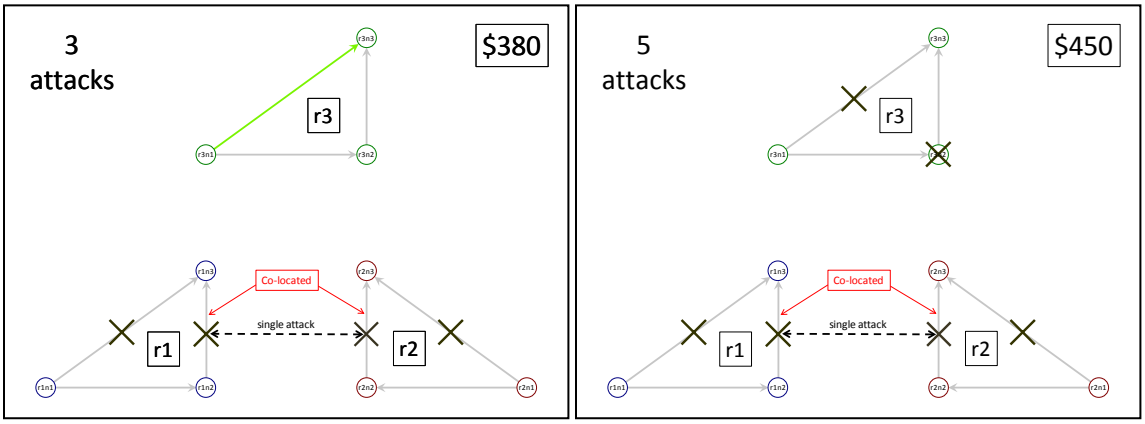

Figure 16. Model Results for a direct dependence between two arcs in separate infrastructures.

An attack on either one interdicts both arcs, effectively disrupting two infrastructures with three attacks (vice four in the base case), and enabling total flow disruption across the collection with five attacks (vice six).

The increases in operating costs for the defender in the presence of attack are shown in Figure 17. In this simple scenario, if the attacker possesses the ability to attack three or more components, the direct dependence results in a cost increase of 5%–15% for the defender. In addition, the attacker can disrupt all flow across the collection with only 5/6 of his original attack resources when compared to the base case. In the presence of direct dependence relationships, treating infrastructures in isolation or as a collection of independent systems can underestimate the potential disruptions.

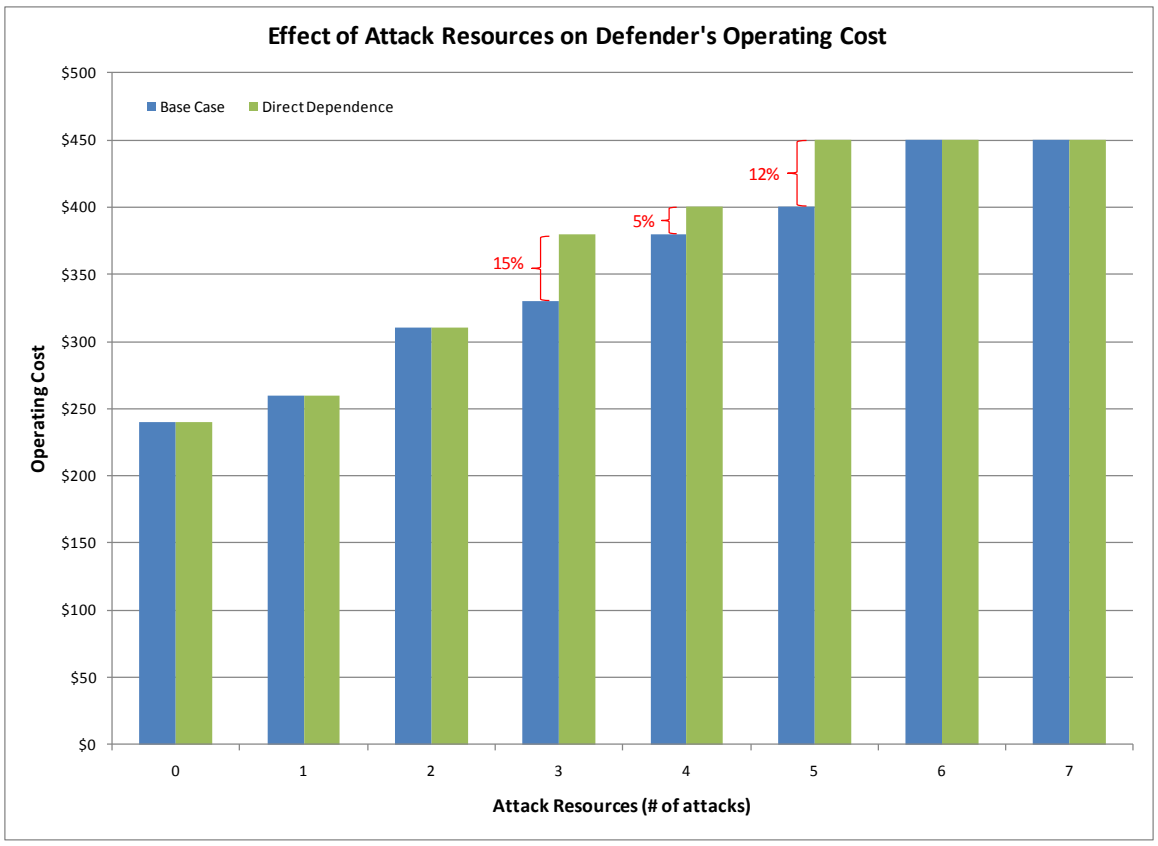

Figure 17. Operating Costs versus Attack Resources.

Comparison of a base case of independent infrastructures against a scenario with a single direct dependence between a pair of arcs in separate infrastructures. The dependence creates higher costs for the defender (15% for 3 attacks) and results in total disruption of the collection of infrastructures with fewer attacks (5 versus 6) through the introduction of a vulnerability.

C. INTERDEPENDENT INFRASTRUCTURES

We now present a different base case for comparing the impact of interdependent infrastructures. Again, consider three independent infrastructures (denoted here as *r1*, *r2*, *r3*), each consisting of three nodes and three arcs and each managed by an individual operator who attempts to satisfy supply and demand at minimum cost. By construction, during normal operation with no interdictions, the low-cost path for each infrastructure now flows through the transshipment node (at a cost of $10 per unit flow), and the direct path from the supply to demand node is more expensive ($11 per unit flow). This collection of infrastructures is shown in Figure 18.

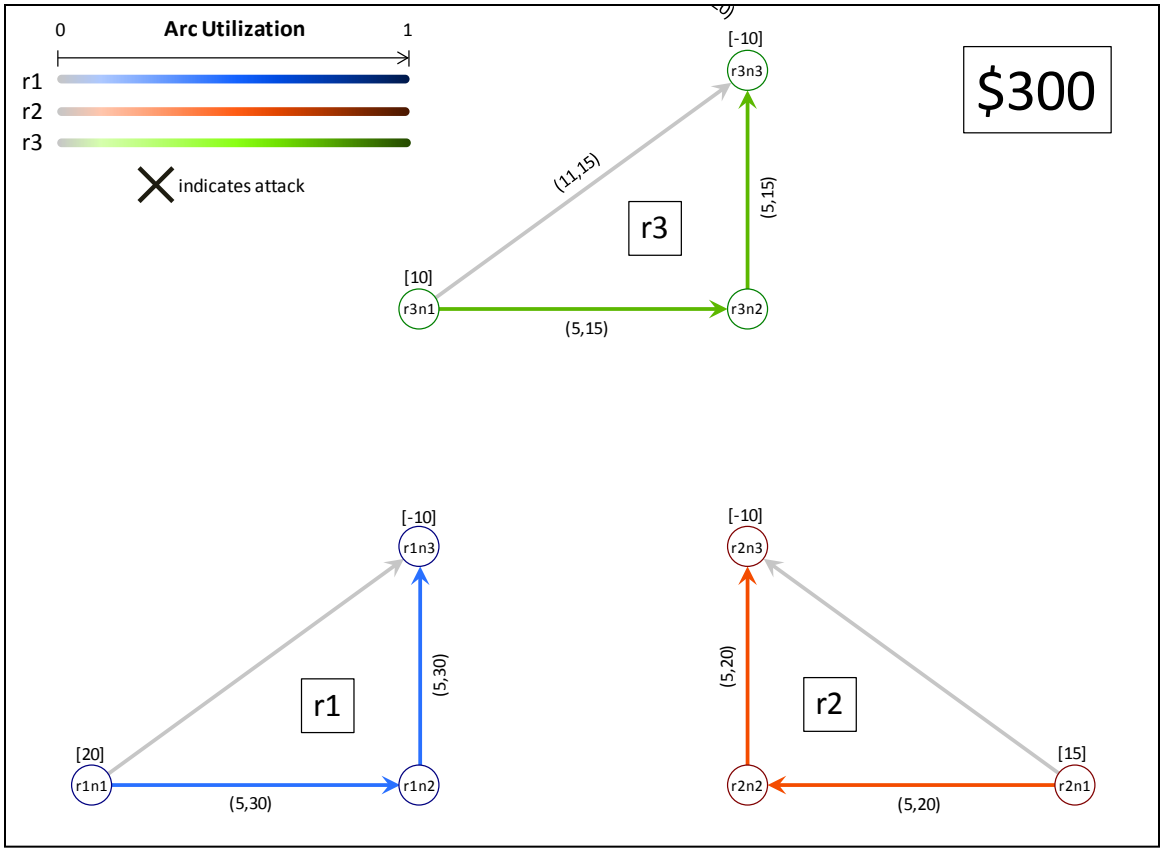

Figure 18. Multiple *independent* infrastructures during normal operation, serving as a base case for *indirect* dependence model.

1. Model Input

Consider a case where all infrastructure costs are in dollars $\left(h^r = 1, \forall r \in R\right)$ and a global manager weights policy costs equally $\left(p^r = 1, \forall r \in R\right)$. All supply and demand nodes are invulnerable to attack; however, we assume that all arcs and transshipment nodes across the collection of infrastructures are vulnerable to attack. We also assume that each infrastructure has sufficient storage capability, so no penalties are charged for excess commodity, but penalty costs for demand shortages ($50/unit flow) are set to encourage commodity flow across interdicted arcs. The per-unit cost to operate across an interdicted arc is $10 for an arc originating or terminating at a transshipment node and $20 for any arc direct between supply and demand nodes (i.e., *r1n1, r1n3*). The cost to operate an interdicted node is now higher, at $25. No direct dependence relationship

exists; interdicted arcs and nodes do not have a direct cost effect on any other arcs or nodes $\left[q_{ijkl} = 0,\ \forall (i,j) \neq (k,l)\right]$. The model input is tabulated in Table 4.

System Data

r	h^r	p^r
r1	1	1
r2	1	1
r3	1	1

Node Data

n	Vuln	b_n	$ePen_n$	$sPen_n$
r1n1		20	0	
r1n2	1			
r1n3		-10		50
r2n1		15	0	
r2n2	1			
r2n3		-10		50
r3n1		10	0	
r3n2	1			
r3n3		-10		50

Arc Data

i	j	Vuln	c_{ij}	u_{ij}
r1n1	r1n2	1	5	30
r1n1	r1n3	1	11	30
r1n2	r1n3	1	5	30
r2n1	r2n2	1	5	20
r2n1	r2n3	1	11	20
r2n2	r2n3	1	5	20
r3n1	r3n2	1	5	15
r3n1	r3n3	1	11	15
r3n2	r3n3	1	5	15

Interdiction Data

i	j	k	l	q_{ijkl}
r1n2	r1n2	r1n2	r1n2	25
r2n2	r2n2	r2n2	r2n2	25
r3n2	r3n2	r3n2	r3n2	25
r1n1	r1n2	r1n1	r1n2	10
r1n1	r1n3	r1n1	r1n3	20
r1n2	r1n3	r1n2	r1n3	10
r2n1	r2n2	r2n1	r2n2	10
r2n1	r2n3	r2n1	r2n3	20
r2n2	r2n3	r2n2	r2n3	10
r3n1	r3n2	r3n1	r3n2	10
r3n1	r3n3	r3n1	r3n3	20
r3n2	r3n3	r3n2	r3n3	10

Table 4. Model Input – Multiple *independent* infrastructures serving as base case for *indirect* dependence scenario.

2. Initial Results

During normal operation of this collection of infrastructures, the optimal commodity flow is across the transshipment node in each infrastructure. Therefore, if the attacker can afford only a single attack, the optimal attack plan is to disrupt the transshipment nodes first (cost of operation on an interdicted node is greater than for an interdicted arc). Because the three infrastructures are identical, attacker selection of an infrastructure is arbitrary. If he can afford two attacks, the optimal attack plan involves interdiction of both the arc between the supply and demand nodes and the transshipment node within a single infrastructure, and the resulting optimal commodity shipment involves flow across the interdicted arc in order to satisfy demand. Maximum cost to the defender results from six attacks, and the attacker does not benefit from further resource increases. Figure 19 highlights the results of this base case.

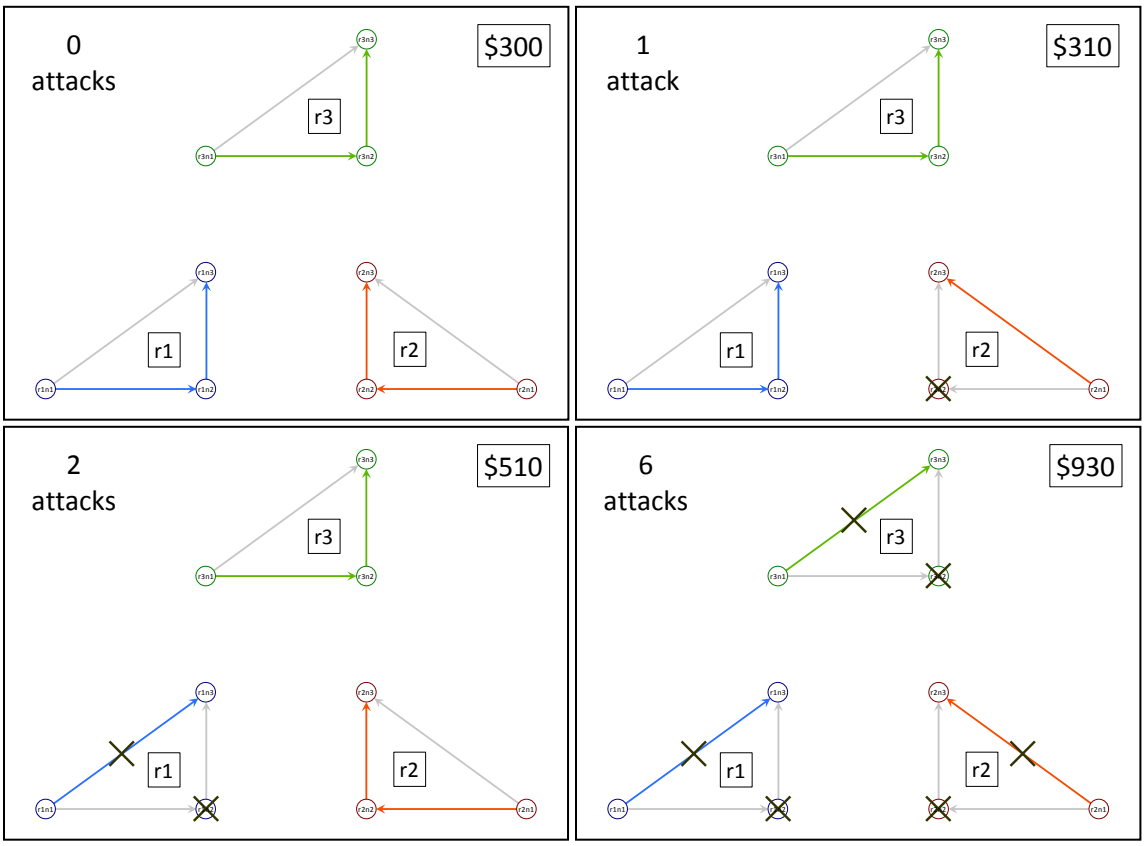

Figure 19. Model Results for multiple *independent* infrastructures.

The minimum-cost path in each infrastructure is through a transshipment node. The optimal single attack targets a transshipment node, requiring selection of the alternate, higher-cost path by the operator. For two attacks, both paths in any infrastructure are interdicted. With all infrastructures identical (including cost conversion factors and policy weights), infrastructure selection for attack is arbitrary. As it is cheaper to operate on the interdicted arc ($20/unit flow) than to operate interdicted nodes ($25/unit flow) or suffer a commodity shortage ($50/unit flow), the optimal flow plan requires shipment across the interdicted arc between the supply and demand nodes. The attacker cannot improve with more than six attacks, when in the best possible commodity flows cross interdicted arcs in all three infrastructures.

3. Indirect Dependence

Consider the case where there are two dependence relationships in this collection of infrastructures.

First, suppose that there is a *shared* dependence between the parent node *r1n3* and two child arcs (*r2n2*, *r2n3*) and (*r3n1*, *r3n2*). Node *r1n3* is capable of supporting both child arcs simultaneously ($max\_supportable_{r1n3}$=2), and each child arc requires five commodity units to operate ($threshold_{[r1n3,(r2n2,r2n3)]}$=$threshold_{[r1n3,(r3n1,r3n2)]}$=5). While child arc (*r2n2*, *r2n3*) requires commodity from only the single parent node ($min\_required_{(r2n2,r2n3)}$=1), a complimentary dependence exists between child arc (*r3n1*, *r3n2*) and parent nodes *r1n3* and *r2n3*. In this case, child arc (*r3n1*, *r3n2*) requires five commodity units from both parent nodes to operate ($min\_required_{(r3n1,\ r3n2)}$=2). Table 5 shows the additional model input while Figure 20 displays the resulting collection of infrastructures.

Dependence Data

n	*i*	*j*	$max\_supportable_{n}$	$min\_required_{ij}$	$threshold_{nij}$
r1n3	r2n2	r2n3	2	1	5
r1n3	r3n1	r3n2	2	2	5
r2n3	r3n1	r3n2	1	2	5

Table 5. Model Input –*indirect* dependence.

Dependence data to support the *shared* and *complementary* dependence relationships. *System*, *Node*, *Arc* and *Interdiction Data* remains as in Table 4.

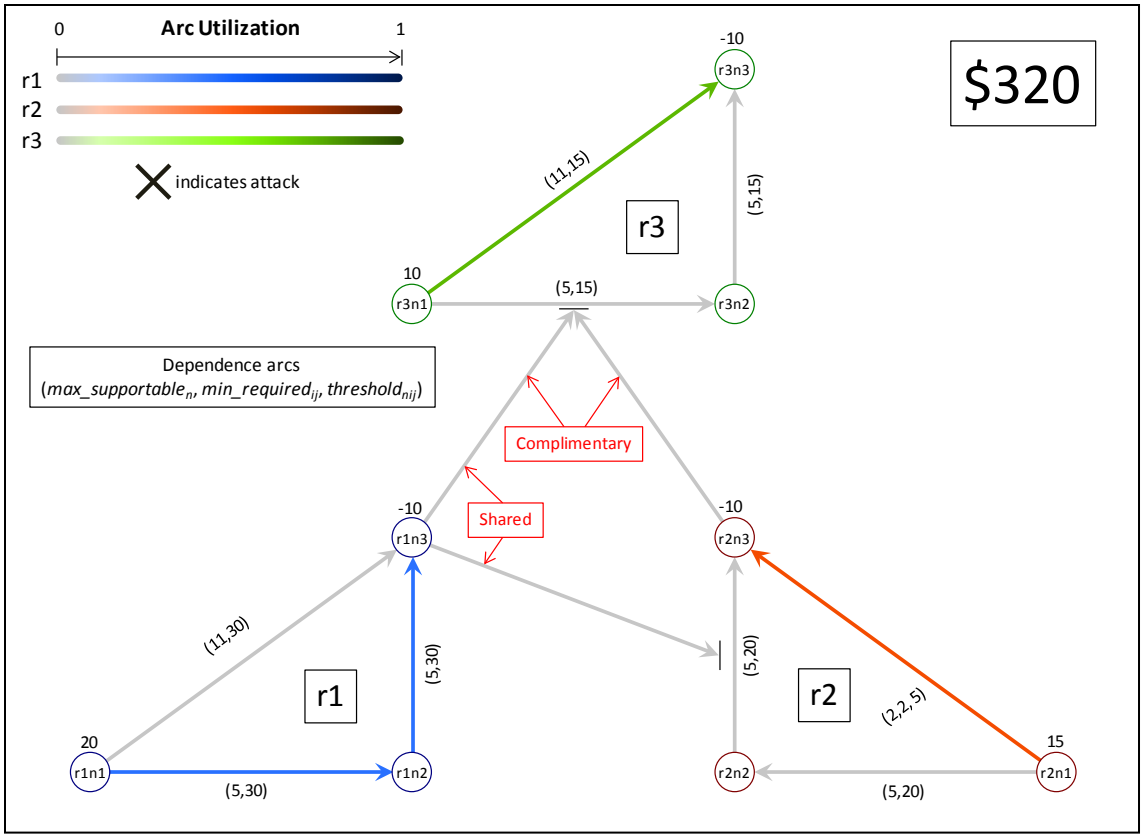

Figure 20. *Indirect* Dependence.

Collection of infrastructures from the base case with a *shared* and *complimentary* dependence added. Each dependence arc is labeled with $max\_supportable_n$, $min\_required_{ij}$, and $threshold_{nij}$, and terminates at a black bar beside the recipient child arc. The optimal flow plan with no interdictions does not use child arcs (*r2n3*,*r2n3*) and (*r3n1*, *r3n2*) due to the additional flow costs required in the parent node infrastructures (*r1* and *r2*) to support these child arcs. This globally optimal solution differs from the local optimal solution for both *r2* and *r3*, which would require flow commodity through the transshipment nodes.

The presence of indirect dependence relationships in this collection changes the optimal actions of both the attacker and defender, as shown in Figure 21. As a result, examination of the optimal attack plans uncover two attacker priorities: interdicting paths that do not contain child arcs in *supported* infrastructures (*r2* or *r3*), and when resources allow, attacking all paths in *supporting* infrastructures (*r1* or *r2*).

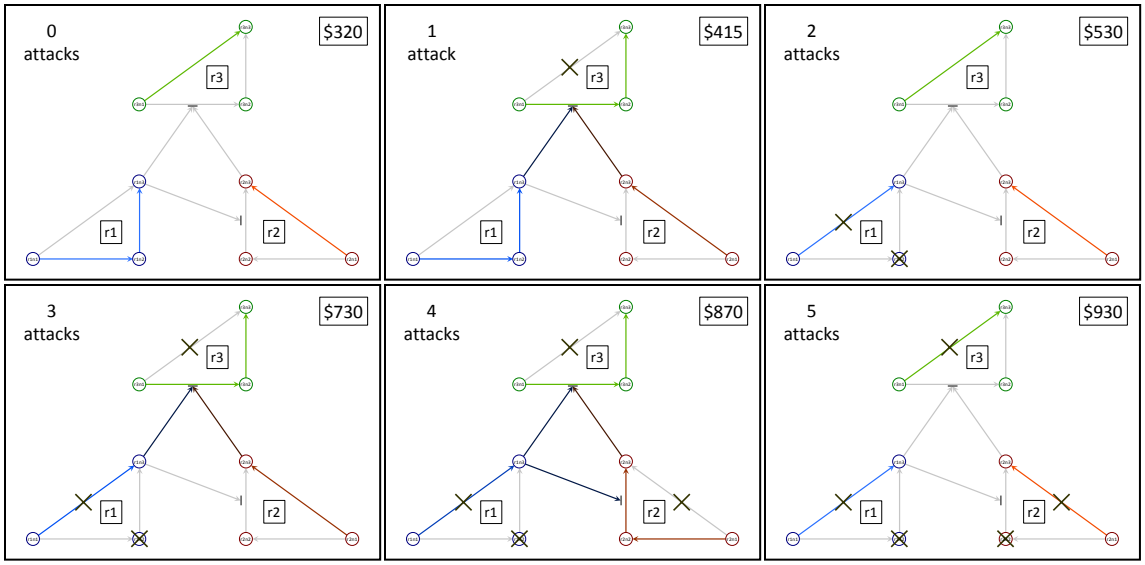

Figure 21. Model Results for *indirect* dependence.

With *indirect* dependence relationships, the minimum-cost paths through *r2* and *r3* are direct from the supply to demand node, to avoid costs associated with using child arcs. The optimal single attack targets arc (*r3n1*, *r3n3*), resulting in use of the most costly dependence; in this case, the *complimentary* relationship from parent nodes *r1n3* and *r2n3* to child arc (*r3n1*, *r3n2*). For two attacks, both paths in *r1* are interdicted, as the parent node in this infrastructure (*r1n3*) potentially supports two separate child arcs, (*r2n2*, *r2n3*) and (*r3n1*, *r3n2*). The defender's optimal response does not make use of either child arc requiring dependence flow. With three attacks, the optimal plan is to interdict both paths in *r1* along with an attack on (*r3n1*, *r3n3*) to force use of the *complimentary* dependence. A fourth attack is placed in *r2* to force use of the *shared* dependence. Five attacks result in shipment across interdicted arcs in all infrastructures with no active dependence relationships. We note that the use of dependence arcs between infrastructures, and the resulting commodity flow and objective function values change significantly with level of attack resources.

For this example, the interdiction of a path that does not contain a child arc in a *supported* infrastructure results in either subsequent shipment across the interdicted arc or commodity flow across the child arc. Because the latter requires use of a supporting dependence relationship, either option leads to an increased cost to the defender. For example, if the attacker can afford only one attack, the optimal attack is against arc (*r3n1*, *r3n3*). The subsequent optimal defender plan requires use of the child arc supported by

the *complementary* dependence, resulting in additional commodity flow costs in both *r1* and *r2* to support the additional demand on the parent nodes *r1n3* and *r2n3*.

Attacks on parent node infrastructures (*r1* and *r2*) further increase the cost to a defender. Consider the situation where the attacker can afford two attacks. The optimal attacker plan is to interdict both paths in *r1*, because the parent node *r1n3* supports two separate dependence relationships. Therefore, *r1* has the highest commodity flow potential and is the optimal infrastructure of the three to interdict.

An intelligent attacker needs no more than five attacks, when he has interdicted both paths in the *supporting* infrastructures (*r1* and *r2*) and also interdicted the single arc in *r3* that does not require use of dependence relationships. Any additional increase in attack resources results in identical defender flow response. The cost to the global manager to operate his collection of infrastructures is higher for every level of interdiction than it is without consideration of indirect dependence relationships, as should be expected. This is graphically displayed in Figure 22.

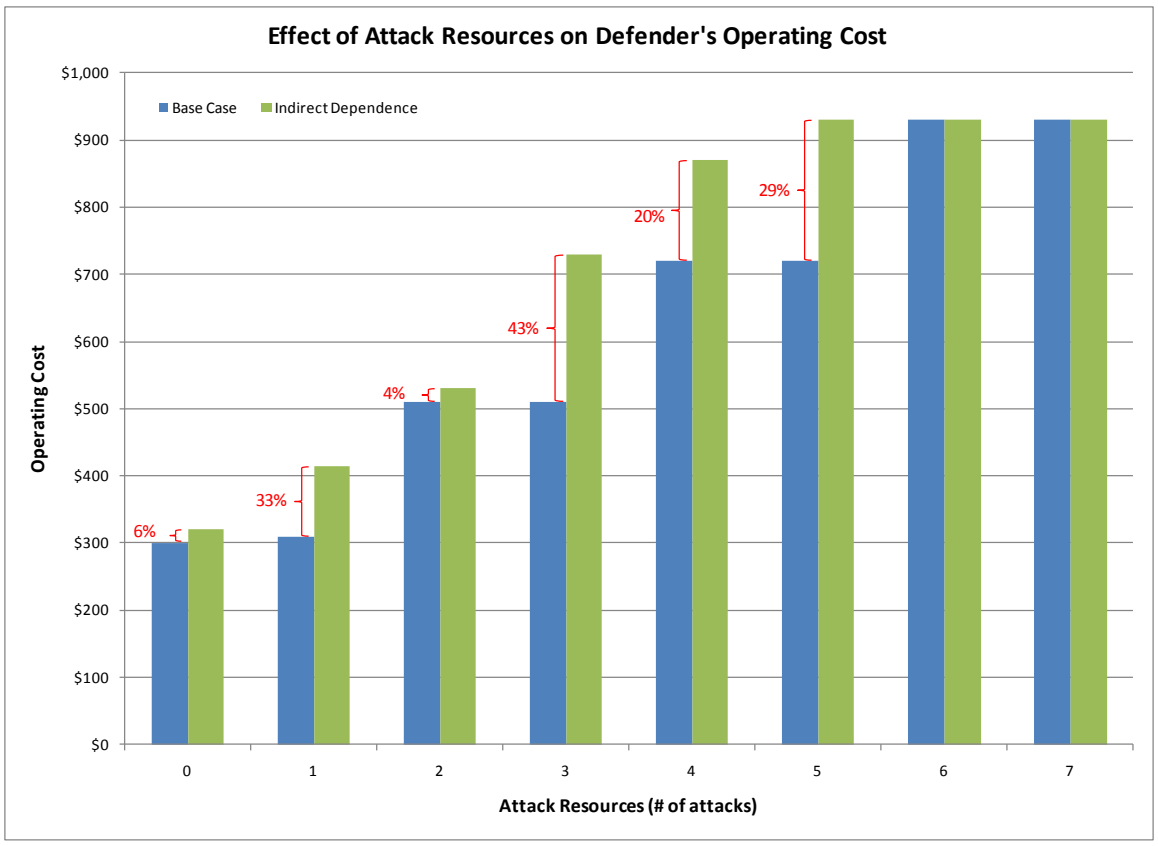

Figure 22. Operating Costs versus Attack Resources.

Comparison of a base case of independent infrastructures against a scenario with two *indirect* dependence relationships. The dependence results in total disruption of the collection of infrastructures with 20% fewer attack resources (five versus six attacks). On average, the global manager's operating cost increases 23% over the base case with no *indirect* dependence, and peaks at 43% for the three-attack scenario.

Using examples like this to examine infrastructure interdependence, we see that with only a small number of interdependent relationships, the minimum-cost operation and worst-case disruptions of infrastructures can become non-intuitive. While there are many more relationships that could be presented, our intent here is to provide simple convincing illustrations, rather than provide an exhaustive collection of examples. We show that adding interdependence relationships creates opportunities for (but does not guarantee) new vulnerabilities. Conversely, finding alternate means of satisfying certain dependence relationships (e.g., *substitute* dependence) might reduce vulnerabilities in a collection of infrastructures, and therefore improve its overall resilience.

IV. CONCLUSIONS AND RECOMMENDATIONS

We conclude by summarizing our work and proposing several ideas for future research on this topic.

A. SUMMARY

This thesis extends the application of attacker-defender models from single to multiple, interdependent infrastructures. We present a general formulation for assessing resilience of a collection of independent infrastructures. We define a *direct*, cost-based dependence and introduce a model to examine such relationships (e.g., *geographic* dependence). Finally, we define six *indirect* component-level dependence relationships: *single-input*, *exclusive-or*, *shared*, *substitute*, *complimentary* and *mutual*; and present a final formulation to assess the resilience of a collection of infrastructures containing both direct and indirect dependence relationships. We present an algorithm based on Benders decomposition to solve this formulation in an efficient manner.

We solve a sequence of simple network flow models and present the worst-case attacks and resulting operator flows for different levels of attacker resources. As our demonstrations show, the assumption that supporting infrastructures are available and invulnerable to attack (as most researchers modeling infrastructures in isolation have done to date) can lead to inaccurate, unjustifiably optimistic assessments of network resilience and can provide operators with a false sense of security. Disruptions can be more costly when infrastructures are interdependent, and the presence of these dependence relationships favors the attacker. We show that locally optimal decisions of a single operator do not always lead to globally optimal behavior within a collection of interdependent infrastructures, necessitating the need for a global decision maker to coordinate such activities at the level of the entire collection of infrastructures.

B. FUTURE WORK

1. Regional Case Study

A natural next step is to apply the techniques in this thesis to a regional case study. Demonstrating the formulation's viability using real-world infrastructures would serve to highlight the extent to which direct and indirect dependence relationships are present within our nation's critical infrastructures. It would also highlight the vulnerabilities we ignore when modeling infrastructures in isolation.

2. Model Refinements

This thesis made several simplifying assumptions that are not realistic and might need to be relaxed in practice. First, our model does not allow for attacks on indirect dependence arcs between infrastructures. Real-world dependence relationships can be vulnerable to attack, and understanding the implications on system resilience is a topic for future research. In addition, we assume for illustrative purposes that an attack on an activity always results in a single, maximum increased cost for a defender to conduct an activity post-interdiction. It may prove advantageous to model levels of attack severity, with the attacker able to choose both the components to interdict and the severity of the attack. Attacker resources can then be defined in terms of a budget, with costs assigned to attack a particular activity at a given level of severity. Each of these model refinements can be made without increasing model complexity.

3. Independent Infrastructure Modeling Techniques

We formulated and solved each example in this thesis as a monolithic model. However, our formulation is separable by design, so that if the linking variables (T, V and W) are fixed, each of the individual infrastructures in the collection can be solved independently, as often happens in the real world. This separation of infrastructures will require more complicated solution techniques. While the master problem (attacker's problem) will remain the same, the defender's problem that serves as the subproblem (**INDIRECT-AD**) will require decomposition to solve. This subproblem will consist of a

global manager's master problem, containing the dependence relationships between infrastructures, and operator subproblems for each individual infrastructure.

4. Additional Dependence Relationships

Although we use minimum-cost network flow models in this thesis for ease of illustration, our main contribution is independent of the particular models used to represent the individual infrastructures. A natural extension of our formulation would be to consider dependence beyond the pure physical relationships defined in this work and cover other dependence classes, such as logical or cyber. A challenge in that domain is how to formulate operator activity $(Y \in \Psi)$ and attacker resource $(X \in \Gamma)$ constraints for each infrastructure within the collection.

5. Extension to Tri-level Defender-Attacker-Defender Models

Other researchers have demonstrated the value of both bi-level and tri-level models for worst-case analysis of infrastructures in isolation (Brown et al. 2006, 2008; Alderson et al. 2011). A natural extension of this thesis work is the implementation of a tri-level model to identify an optimal defensive plan for the collection of infrastructures. Much like the extension to separable infrastructures, the addition of the third level, the defender's preparation problem, is not a trivial matter and will create new computational complexities.

C. FINAL THOUGHTS

Accidents such as the train derailment in Baltimore's Howard Street Tunnel and terrorist attacks such as 9/11 have served to highlight the interdependencies present among our nation's critical infrastructures, as well as the negative impacts that can result. We provide our formulations as a means of representing these dependence relationships in operational-level, game-theoretic models to uncover resulting vulnerabilities and more accurately assess resilience for collections of infrastructures.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Ahuja, R. K., Magnanti, T. L., & Orlin, J. B. (1993). *Network flows.* New Jersey: Prentice Hall.

Alderson, D., Brown, G., Carlyle, M., & Wood, K. (2011). Solving defender-attacker-defender models for infrastructure defense. *Operations Research, Computing and Homeland Defense,* Wood and R.F. Dell, editors, INFORMS, Hanover, MD, pp. 28–49.

Benders, J. (1962). Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik 4,* pp. 238–252.

Bernstein, A., Bienstock, D., Hay, D., Uzunoglu, M., & Zussman, G. (2011). Power grid vulnerability to geographically correlated failures - Numerical Evaluation. *Columbia University, Electrical Engineering Technical Report #2011-05-06,* pp. 1–9.

Brown, G., Carlyle, M., Diehl, D., Kline, J., & Wood, K. (2005). A two-sided optimization for theater ballistic missile defense. *Operations Research, 53* (5), pp. 745–763.

Brown, G., Carlyle, M., Salmerón, J., & Wood, K. (2005). Analyzing the vulnerability of critical infrastructure to attack and planning defenses. *Tutorials in Operations Research: Emerging Theory, Methods, and Applications,* pp. 102–123.

Brown, G., Carlyle, M., Salmerón, J., & Wood, K. (2006). Defending critical infrastructure. *Interfaces, 36*(6), pp. 530–544.

Brown, G., Carlyle, M., & Wood, K. (2008). Optimizing Department of Homeland Security defense investments: Applying defender-attacker (-defender) optimization to terror risk assessment and mitigation. Appendix E in *Department of Homeland Security Bioterrorism Risk Assessment: A Call for Change*, National Research Council report, National Academies Press, Washington, DC.

Brown, K. A. (2006). *Critical path: A brief history of critical infrastructure protection in the United States*. Arlington, VA: George Mason University Press.

Chou, C.-C., & Tseng, S.-M. (2010). Collection and analysis of critical infrastructure interdependency relationships. *Journal of Computing in Civil Engineering, 24*(6), pp. 539–547.

Davidson, P. (2008, August 12). *U.S. power grid in better shape 5 years after blackout.* Retrieved May 29, 2011, from USA Today website: http://www.usatoday.com/money/industries/energy/2008-08-12-blackout-power-outage_N.htm

Executive Order No. 13010, 3 C.F.R. 37347 (1996).

Grubesic, T. H., & Murray, A. T. (2006). Vital nodes, interconnected infrastructures, and the geographies of network survivability. *96*(1), pp. 64–83.

Kennedy, K. T., Deckro, R. F., Chrissis, J. W., & Wiley, V. D. (2009). On modeling and analyzing multi-layered networks. *Military Operations Research, 14*(3), pp. 53–66.

Law, A. M., & Kelton, W. D. (2000). *Simulation modeling and analysis.* Boston: McGraw-Hill Companies.

Lee, I. E., Mitchell, J. E., & Wallace, W. A. (2004). Assessing vulnerability of proposed designs for interdependent infrastructure systems. *Proceedings of the 37th Hawaii International Conference on System Sciences,* pp. 1–8.

Lee, I. E., Mitchell, J. E., & Wallace, W. A. (2007). Restoration of services in interdependent infrastructure systems: A network flows approach. *IEEE Transactions on Systems, Man, and Cybernetics part C, 37* (6), pp. 1303–1317.

MyFox Memphis. (2010, September 08). *Lane closures on I-40 bridge begin Tuesday*. Retrieved June 12, 2011, from MyFox Memphis website: http://www.myfoxmemphis.com/dpp/news/local/090210-lane-closures-on-i-40-bridge-begin-september-7

National Infrastructure Advisory Council (NIAC). (2009). *Critical infrastructure resilience final report and recommendations.* Retrieved from Department of Homeland Security website: http://www.dhs.gov/xlibrary/assets/niac/niac_critical_infrastructure_resilience.pdf

President's Commission on Critical Infrastructure Protection (PCCIP). (1997a). *Critical foundations: Protecting America's infrastructures.* Retrieved from Federation of American Scientists website: http://www.fas.org/sgp/library/pccip.pdf

President's Commission on Critical Infrastructure Protection (PCCIP). (1997b). *Critical foundations: Thinking differently.* Retrieved from The Information Warfare Site website: http://www.iwar.org.uk/cip/resources/pccip/summary.pdf

Rinaldi, S., Peerenboom, J., & Kelly, T. (2001). Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE Control Systems Magazine*, December 2001, pp. 11–25.

Rinaldi, S. M. (2004). Modeling and simulating critical infrastructures and their interdependencies. *Proceedings of the 37th Hawaii International Conference on System Sciences*, pp. 1–8.

Robert, B., & Marabito, L. (2010). An approach to identifying geographic interdependencies among critical infrastructures. *Int. J. Critical Infrastructures , 6* (1), pp. 17-30.

Salmerón, J., Wood, K., & Baldick, R. (2004). Analysis of electric grid security under terrorist threat. *IEEE Transactions on Power Systems*, *19*(2), pp. 905–912.

Salmerón, J., Wood, K., & Baldick, R. (2009). Worst-case interdiction analysis of large-scale electric power grids. *IEEE Transactions on Power Systems, 24*(1), pp. 96–104.

U.S. Department of Homeland Security (DHS). (2007). *National strategy for Homeland Security.* Retrieved from Department of Homeland Security website: http://www.dhs.gov/xlibrary/assets/nat_strat_homelandsecurity_2007.pdf

U.S. Department of Homeland Security (DHS). (2009). *National infrastructure protection plan: Partnering to enhance protection and resiliency.* Retrieved from Department of Homeland Security website: http://www.dhs.gov/xlibrary/assets/NIPP_Plan.pdf

U.S. Department of Transportation (DoT) ITS Joint Program Office. (2002). *Effects of catastrophic events on transportation system management and operations.* Retrieved from Research and Innovative Technology Administration National Transportation Library: http://ntl.bts.gov/lib/jpodocs/repts_te/13754_files/13754.pdf

Wallace, W. A., Mendonca, D. M., Mitchell, J. E., & Chow, J. H. (2003). Managing disruptions to critical interdependent infrastructures in the context of the 2001 World Trade Center attack. *Impacts of and Human Response to the September 11, 2001 Disasters: What Research Tells Us*, pp. 165–198.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Ryan Timmons
   Research Scientist, O/A011S
   Sunnyvale, California

4. Professor David L. Alderson
   Naval Postgraduate School
   Monterey, California

5. Professor Gerald G. Brown
   Naval Postgraduate School
   Monterey, California

6. Professor W. Matthew Carlyle
   Naval Postgraduate School
   Monterey, California